

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:00:58 ; Search time 50.2093 Seconds
(without alignments)
213.658 Million cell updates/sec

Title: US-10-010-667a-2_COPY_185_218

Perfect score: 192

Sequence: 1 RRSYRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mbc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	192	100.0	338	6 Q9GL50	Q9GL50 sus scrofa
2	187	97.4	339	11 Q924J9	Q924J9 mus musculus
3	187	97.4	339	11 Q9CWR7	Q9CWR7 mus musculus
4	139	72.4	339	11 Q924Z2	Q924Z2 mus musculus
5	110	57.3	489	11 Q8BWB6	Q8BWB6 mus musculus
6	109	56.8	488	11 Q99P41	Q99P41 rattus norv
7	109	56.8	488	11 Q8CI59	Q8CI59 mus musculus
8	109	56.8	514	11 Q924Z1	Q924Z1 mus musculus
9	109	56.8	526	11 Q80ZP3	Q80ZP3 mus musculus
10	107	55.7	488	4 Q9NVB5	Q9NVB5 homo sapien
11	107	55.7	488	4 Q86SF6	Q86SF6 homo sapien
12	107	55.7	490	4 Q8NFT2	Q8NFT2 homo sapien
13	107	55.7	490	4 Q8IUE7	Q8IUE7 homo sapien
14	107	55.7	498	4 Q7Z389	Q7Z389 homo sapien
15	105	54.7	488	11 Q8CSF0	Q8CSF0 mus musculus
16	96.5	50.3	456	4 Q8TDP3	Q8TDP3 homo sapien

SUMMARIES

17	96.5	50.3	487	4	Q8TF03	Q8tf03 homo sapien
18	96.5	50.3	487	4	Q8NEW6	Q8new6 homo sapien
19	62.5	32.6	480	16	Q67595	Q67595 aquifex ae
20	62	32.3	366	16	Q8A6D0	Q8a6d0 bacteroides
21	60	31.2	926	4	Q8TE49	Q8te49 homo sapien
22	60	31.2	926	11	Q8RS54	Q8rs54 mus musculu
23	58.5	30.5	1027	5	Q95QC7	Q95qc7 caenorhabdi
24	58	30.2	213	16	Q813H4	Q813h4 bacillus ce
25	57.5	29.9	556	11	Q8CFS0	Q8cfs0 mus musculu
26	57.5	29.9	858	4	Q9NQ53	Q9nq53 homo sapien
27	56	29.2	150	12	Q9YVJ4	Q9yv74 melanoplus
28	56	29.2	230	2	Q9RFP3	Q9rfp3 mycoplasma
29	55.5	28.9	502	2	Q48904	Q48904 microcystis
30	55	28.6	81	16	Q8IMC0	Q8imc0 bacillus an
31	55	28.6	1379	5	Q817F4	Q817f4 dictyosteli
32	54	28.1	81	16	Q812U5	Q812u5 bacillus ce
33	53.5	27.9	427	4	Q8MWA7	Q8mwa7 homo sapien
34	53.5	27.9	831	16	Q8D8X5	Q8d8x5 vibrio vuln
35	53	27.6	152	17	Q30299	Q30299 archaeoglob
36	53	27.6	307	16	Q8EAP3	Q8eap3 shewanella
37	53	27.6	336	2	Q87519	Q87519 escherichia
38	53	27.6	391	16	Q8FEB8	Q8feb8 escherichia
39	53	27.6	429	17	Q8TFE0	Q8tf0 methanosarc
40	53	27.6	536	16	Q89Q05	Q89q05 bradyrhizob
41	52.5	27.3	533	16	Q8ZMD6	Q8zmd6 streptomyce
42	52.5	27.3	685	4	Q81WK5	Q81wk5 homo sapien
43	52	27.1	132	15	Q91WQ0	Q91wg0 human immun
44	52	27.1	252	4	Q8N2A8	Q8n2a8 homo sapien
45	52	27.1	252	4	Q8N5Y1	Q8n5y1 homo sapien

ALIGNMENTS

RESULT 1

Q9GL50 PRELIMINARY; PRT; 338 AA.

AC Q9GL50; TREMBLrel. 16, Created
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Six transmembrane endothelial antigen of PAEC.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Nagasaka T., Boulday G., Coupel S., Coulon F., Tesson L.,
RA Heslan J.-M., Soullou J.-P., Charreau B.;
RT "Differential gene expression in endothelial cells during TNF-alpha-
RT and LPS-mediated activation."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF319659; AAG33868.1; --
DR GO: GO:0016021; C:integral to membrane; IEA.
KW Transmembrane.
SQ SEQUENCE 338 AA; 39918 MW; ED490E86E067A32B CRC64;

Query Match 100.0%; Score 192; DB 6; Length 338;
Best Local Similarity 100.0%; Pred. No. 7.9e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRSYRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34

Db 184 RRSYRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 217

RESULT 2

Q924J9 PRELIMINARY; PRT; 339 AA.

ID Q924J9

AC Q924J9

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

```
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Six-transmembrane epithelial antigen of the prostate.
GN STEAP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
EN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic stem cells;
RX MEDLINE=21371909; PubMed=11479226;
RA Yang D., Holt G.E., Velders M.P., Kwon E.D., Kast W.M.;
RT "Murine six-transmembrane epithelial antigen of the prostate, prostate
RT stem cell antigen, and prostate-specific membrane antigen: prostate-
RT specific cell-surface antigens highly expressed in prostate cancer of
RT transgenic adenocarcinoma mouse prostate mice.";
RL Cancer Res. 61:5857-5860(2001).
DR EMBL; AF297098; AAK83126.1; -.
DR MGD; MGI:1917608; Steap.
DR GO; GO:0016021; C:integral to membrane; IEA.
KW Transmembrane.
SQ SEQUENCE 339 AA; 39318 MW; 4B26A71FF59E84F CRC64;

Query Match 97.4%; Score 187; DB 11; Length 339;
Best Local Similarity 94.1%; Pred. No. 3.5e-16;
Matches 32; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRSRYKLLNWAYQVQONKEDAWIEHDVWRMEI 34
|||||:|||||:|||||:|||||:|||||
Db 185 RRSRYKLLNWAYQVQONKEDAWVEHDVWRMEI 218
|||||:|||||:|||||:|||||:|||||

RESULT 3
Q9CWR7 ID Q9CWR7 PRELIMINARY; PRT; 339 AA.
AC Q9CWR7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE 241007B19RIK protein.
GN STEAP OR 241007B19RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
EN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic stem cells;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itch M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauber J.F., Suzuki R., Tonita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010437; BAB26938.1; -.
DR MGD; MGI:1917608; Steap.
SQ SEQUENCE 339 AA; 39264 MW; 3F7AB9C7520F0968 CRC64;
```

```
Query Match 97.4%; Score 187; DB 11; Length 339;
Best Local Similarity 94.1%; Pred. No. 3.5e-16;
Matches 32; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRSRYKLLNWAYQVQONKEDAWIEHDVWRMEI 34
|||||:|||||:|||||:|||||:|||||
Db 185 RRSRYKLLNWAYQVQONKEDAWVEHDVWRMEI 218
|||||:|||||:|||||:|||||:|||||

RESULT 4
Q924Z2 ID Q924Z2 PRELIMINARY; PRT; 339 AA.
AC Q924Z2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dudulin.
GN STEAP OR 1010001D0IRIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
EN [1]
RP SEQUENCE FROM N.A.
RA Seru V., Manivet P., Lamblin D., Vaubourdolle M., Kellermann O.,
RA Loric S.;
RT "Prostate and non-prostate expression of dudulin, the mouse ortholog
RT of human STEAP.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY029584; AAK50537.1; -.
DR MGD; MGI:1917608; Steap.
SQ SEQUENCE 339 AA; 39109 MW; 32A2C29F2E33BD0 CRC64;

Query Match 72.4%; Score 139; DB 11; Length 339;
Best Local Similarity 73.5%; Pred. No. 5.8e-10;
Matches 25; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RRSRYKLLNWAYQVQONKEDAWIEHDVWRMEI 34
|||||:|||||:|||||:|||||:|||||
Db 185 RRSRYDLNLAVKQVLANKEDAWVEHDVWRMEI 218
|||||:|||||:|||||:|||||:|||||

RESULT 5
Q8BW56 ID Q8BW56 PRELIMINARY; PRT; 489 AA.
AC Q8BW56;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Weakly similar to tumor suppressor PHYE.
GN 4921538B17RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
EN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK052981; BAC35230.1; -.
DR MGD; MGI:1921301; 4921538B17RIK.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred F420.
DR Pfam; PF03807; F420_oxidored; 1.
SQ SEQUENCE 489 AA; 55760 MW; 98CD63D59DDDF24C CRC64;

Query Match 57.3%; Score 110; DB 11; Length 489;
Best Local Similarity 61.8%; Pred. No. 4.8e-06;
```


DR	GO; GO:0006118; P:electron transport; IEA.
DR	InterPro; IPR004455; NADPoxred.F420.
DR	Pfam; PF03807; F420 oxidored; I.
SQ	SEQUENCE 488 AA; -54749 MW; 9A08D9C9C0CF83F4 CRC64;

Query Match 56.8%; Score 109; DB 11; Length 488;
Best Local Similarity 58.8%; Pred.No.6.5e-06;
Matches 20; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY	1 RRSVRYKLANWAVQQOONKEADAWIHDVVRMEI 34
Dd	: : : : : : 326 RRSHRDVLNLAVKQLANKSLRWEEVWRMEI 359

RESULT 8	
Q924Z1	PRELIMINARY; PRT; 514 AA.
ID Q924Z1	AC Q924ZI AC Q924ZI
DT 01-DEC-2001	(TrEMBLrel. 19, Created)
DT 01-DEC-2001	(TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003	(TrEMBLrel. 24, Last annotation update)
DE Dudulin 2,	
GN TSAP6 OR i010001D0IRIK.	
OS Mus musculus (Mouse).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
NCBI_TaxID=10090;	
[1]	
RN SEQUENCE FROM N.A.	
RP Serru V., Lamblin D., Lenoir C., Manivet P., Vaubourdiolle M., Kellermann O., Lorix S.; "Molecular cloning and expression of mouse dudulin 2."; Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.	
RL EMBL; AY029586; AAK50539.1; -. ENMET; MGII:1915678; Tsap6.	
DR MGD; MGII:1915678; Tsap6.	
DR GO; GO:0006118; P:electron transport; IEA.	
DR InterPro; IPR004455; NADPoxred.F420.	
DR Pfam; PF03807; F420 oxidored; I.	
SQ SEQUENCE 514 AA; -57268 MW; 3398B6C288AECE0E2 CRC64;	

Query Match 56.8%; Score 109; DB 11; Length 514;
Best Local Similarity 58.8%; Pred.No.6.9e-06;
Matches 20; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY	1 RRSVRYKLANWAVQQOONKEADAWIHDVVRMEI 34
Dd	: : : : : : 326 RRSHRDVLNLAVKQLANKSLRWEEVWRMEI 359

RESULT 9	
Q80ZF3	PRELIMINARY; PRT; 526 AA.
ID Q80ZF3	AC Q80ZF3 AC Q80ZF3
DT 01-JUN-2003	(TrEMBLrel. 24, Created)
DT 01-JUN-2003	(TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003	(TrEMBLrel. 25, Last annotation update)
DE Tsap6.	
OS Mus musculus (Mouse).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
NCBI_TaxID=10090;	
[1]	
RN SEQUENCE FROM N.A.	
RC STRAIN=DSMZ AC331; TISSUE=Bone marrow;	
RR MEDLINE=22506415; PubMed=12606722;	
RA Passer B.J., Nancy-Portebois V., Amzallag N., Prieur S., Cans C., Roborel de Climens A., Fiucci G., Bouvard V., Tuynder M., Susini L., Morchoisne S.P., Cribie V., Lespagnol A., Dausset J., Oren M., Anson R., Tielerman A.; "The p53-inducible TSAP6 gene product regulates apoptosis and the cell cycle and interacts with Nix and the Myt1 kinase." Proc. Natl. Acad. Sci. U.S.A. 100:2284-2289(2003). ENBL; AY214462; AAC038239.1; -. DR	

DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR004455; NADPoxred_F420.
 DR Pfam; PF03807; F420 oxidored; 1.
 SQ SEQUENCE 526 AA; 58530 MW; 6306CD717E25200A CRC64;

Query Match 56.8%; Score 109; DB 11; Length 526;
 Best Local Similarity 58.8%; Pred. No. 7e-06;
 Matches 20; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 RRSRYRKLWYQQVQOONKEDAWIEHDVWRMEI 34
 |||::|||::|||::|||::|||::|||::|||
 Db 364 RRSHRYDLNVLAVKQVLANKSLRWEEVWRMEI 359

RESULT 10
 QNVBS PRELIMINARY; PRT; 488 AA.

ID QNVBS
 AC QNVBS;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein FLJ10829 (Dudulin 2).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

Rf Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
 RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
 RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
 RT "MEDO human cDNA sequencing project.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
 [2]

RP SEQUENCE FROM N.A.
 RA Serri V., Manivet P., Lenoir C., Eschwege P., Lamblin D.,
 RA Vaubourville M., Kellermann O., Loric S.;
 RT "Dudulin 2, a new tumor antigen expressed in various human tumors.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
 DR ENBL; AK001691; BAA91839.1; -.
 DR ENBL; AY029585; AAK50538.1; -.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR004455; NADPoxred_F420.
 DR Pfam; PF03807; F420 oxidored; 1.
 DR PROSITE; PS00290; IG_MHC; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 488 AA; 54616 MW; BC0BCA49335AAD6 CRC64;

Query Match 55.7%; Score 107; DB 4; Length 488;
 Best Local Similarity 55.9%; Pred. No. 1.2e-05;
 Matches 19; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 1 RRSRYRKLWYQQVQOONKEDAWIEHDVWRMEI 34
 |||::|||::|||::|||::|||::|||::|||
 Db 326 RRAHRYDLNVLAVKQVLANKSLRWEEVWRMEI 359

RESULT 11
 Q86SF6 PRELIMINARY; PRT; 488 AA.

ID Q86SF6
 AC Q86SF6;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Similar to dudulin 2 (TSAP6).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

Rf Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
 RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
 RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
 RT "MEDO human cDNA sequencing project.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
 [2]

RP SEQUENCE FROM N.A.
 RA Serri V., Manivet P., Lenoir C., Eschwege P., Lamblin D.,
 RA Vaubourville M., Kellermann O., Loric S.;
 RT "Dudulin 2, a new tumor antigen expressed in various human tumors.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
 DR ENBL; AK001691; BAA91839.1; -.
 DR ENBL; AY029585; AAK50538.1; -.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR004455; NADPoxred_F420.
 DR Pfam; PF03807; F420 oxidored; 1.
 DR PROSITE; PS00290; IG_MHC; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 488 AA; 54616 MW; BC0BCA49335AAD6 CRC64;

Query Match 55.7%; Score 107; DB 4; Length 488;
 Best Local Similarity 55.9%; Pred. No. 1.2e-05;
 Matches 19; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 1 RRSRYRKLWYQQVQOONKEDAWIEHDVWRMEI 34
 |||::|||::|||::|||::|||::|||::|||
 Db 326 RRAHRYDLNVLAVKQVLANKSLRWEEVWRMEI 359

RESULT 12
 Q8NF2 PRELIMINARY; PRT; 490 AA.

ID Q8NF2
 AC Q8NF2;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Six-transmembrane epithelial antigen of prostate 2.
 DR STEAP2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

Rf Poraka K.P., Helenius M.A., Visakorpi T.;
 RA "Cloning and characterization of a novel six-transmembrane protein
 STEAP2, down-regulated in androgen-independent prostate cancer.";
 RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
 DR ENBL; AF455138; AAN04080.1; -.
 DR GeneW; HGNC:17885; STEAP2.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR004455; NADPoxred_F420.
 DR Pfam; PF03807; F420 oxidored; 1.
 KW Transmembrane.
 SQ SEQUENCE 490 AA; 55961 MW; F2E9C30CDACCEFF81 CRC64;

Query Match 55.7%; Score 107; DB 4; Length 490;
 Best Local Similarity 58.8%; Pred. No. 1.2e-05;
 Matches 20; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 RRSRYRKLWYQQVQOONKEDAWIEHDVWRMEI 34
 |||::|||::|||::|||::|||::|||::|||
 Db 326 RRSERYFLNWAYQQVQHANIENSWNEEVWRMEI 359

RESULT 13
 Q8IU67 PRELIMINARY; PRT; 490 AA.

ID Q8IU67
 AC Q8IU67;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 13:52:02 ; Search time 10.6744 Seconds
(without alignments)

165.853 Million cell updates/sec

Title: US-10-010-667a-2_COPY_185_218

Perfect score: 192

Sequence: 1 RRSRYKLINWAYQQQNKEDAWIEHDVWRMEI 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	192	100.0	339	1 STEA_HUMAN	Q9u8e8 homo sapien
2	58.5	30.5	1418	1 STEA_CAEEL	P34641 caenorhabdi
3	58	30.2	900	1 GGAB_BACSU	P46918 bacillus su
4	54.5	28.4	636	1 SYT_WIGER	Q8d3c0 wigleswort
5	54	28.1	630	1 Y4BJ_RHISN	P55377 rhizobium s
6	52	27.1	79	1 YOKC_BACSU	P54566 bacillus su
7	52	27.1	1813	1 UN13_CAEEL	P27715 caenorhabdi
8	51	26.6	495	1 GLPK_SYNY3	P74260 synecocyst
9	51	26.6	1017	1 ACAD_ARATH	Q911k7 arabidopsis
10	50.5	26.3	313	1 TRUB_PROMA	P59880 prochloroco
11	50	26.0	784	1 ALP4_SCHPO	Q9y705 schizosacch
12	50	26.0	1574	1 SYJ1_RAT	Q62910 rattus norv
13	50	26.0	5430	1 MACF_HUMAN	Q9upn3 homo sapien
14	50	26.0	5938	1 MAC4_HUMAN	Q96pk2 homo sapien
15	49.5	25.8	95	1 YQ04_BACAN	Q9rn28 bacillus an
16	49.5	25.8	274	1 DAPD_BUCAI	P57323 buchnera ap
17	49.5	25.8	350	1 YC73_METJA	Q58669 methanococ
18	49.5	25.8	787	1 OKAA_CHLTR	Q84253 chlamydia t
19	49.5	25.8	940	1 SYV_CHLPN	Q92987 chlamydia p
20	49	25.5	257	1 TRUA_XYLPA	Q9pd66 xyliella fas
21	49	25.5	257	1 TRUA_XYLFT	Q87d51 xyliella fas
22	49	25.5	500	1 GLPK_ANASP	Q8y05 anabaena sp
23	49	25.5	512	1 Y4WA_RHISN	P55679 rhizobium s
24	49	25.5	837	1 NCM2_HUMAN	O15394 homo sapien
25	49	25.5	837	1 NCM2_MOUSE	O35136 mus musculu
26	49	25.5	1324	1 SYJ1_BOVIN	O18964 bos taurus
27	49	25.5	1575	1 SYJ1_HUMAN	O43426 homo sapien
28	48.5	25.3	266	1 IF2A_SULSO	Q97z79 sulfobolus
29	48.5	25.3	339	1 MO2L_CAEEL	Q9tzm2 caenorhabdi
30	48	25.0	286	1 YAY8_SCHPO	O10216 schizosacch
31	48	25.0	405	1 ASSY_PSEPK	P59604 pseudomonas
32	48	25.0	466	1 SYN_SHEON	Q9eez1 shewanella
33	48	25.0	620	1 SYR_BIFLO	Q8g4v2 bifidobacte

RESULT 1

STEA_HUMAN STANDARD; PRT; 339 AA.

AC Q9UHE8; O95034;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Six transmembrane epithelial antigen of prostate.
GN STEAP OR STEAP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20056277; PubMed=10588738;
RA Hubert R.S., Vivanco I., Chen E., Rastegar S., Leong K.,
RA Mitchell S.C., Madraswala R., Zhou Y., Kuo J., Raitano A.B.,
RA Jakobovits A., Safran D.C., Afar D.E.H.;
RT "STEAP: a prostate-specific cell-surface antigen highly expressed in human prostate tumors.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:14523-14528(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Abu-Threideh J., Stoneking T., Langston Y., Maupin R.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22398257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udgin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Marra M.A.;
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: Highly expressed in prostate tumors.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

ALIGNMENTS

34	48	25.0	864	1	AGLU_MUCJA
35	48	25.0	942	1	ENV_CAEVG
36	48	25.0	1028	1	FTSK_VIBPA
37	47.5	24.7	561	1	O16G_BACSU
38	47.5	24.7	796	1	DHG_ECOLI
39	47.5	24.7	808	1	DHG_GLUOX
40	47.5	24.7	1103	1	CYGF_BOVIN
41	47.5	24.7	1122	1	VAB1_CAEEL
42	47	24.5	245	1	TRMD_CAUCR
43	47	24.5	586	1	VATA_HALVO
44	47	24.5	626	1	PEPX_RHOBA
45	47	24.5	675	1	VP85_YEAST

Q92442 mucor javan
P31627 caprine arc
Q87dp4 vibrio para
O06994 bacillus su
P15977 escherichia
P27175 gluconobact
O02740 bos taurus
O61460 caenorhabdi
Q9abm9 caulobacter
Q48332 halobacteri
P59825 rhodopirell
Q92331 saccharomyc

Db 296 RYRRELFGWEYR-----AKLEPEIWR 318

RESULT 6

YQKC_BACSU

ID YQKC_BACSU STANDARD; PRT; 79 AA.

AC P54566;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Hypothetical protein yqkC.

GN YQKC OR BSU23650.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168 / JH642;

RX MEDLINE=97124195; PubMed=8969508;

RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,

RA Kobayashi Y.;

RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of

RT the Bacillus subtilis genome containing the skin element and many

RT sporulation genes.";

RL Microbiology 142:3103-3111(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,

RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,

RA Borris R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,

RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,

RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,

RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,

RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,

RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,

RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,

RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,

RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,

RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,

RA Kobayashi Y., Koetter P., Konigstein G., Krohn S., Kumano M.,

RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,

RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,

RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,

RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,

RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,

RA Rieger M., Rivolta C., Rochelle E., Roche B., Rose M., Sadale Y.,

RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,

RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,

RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,

RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,

RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,

RA Viari A., Wambut R., Wedler E., Wedler H., Weitzenecker T.,

RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,

RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;

RT "The complete genome sequence of the Gram-positive bacterium Bacillus

RT subtilis.";

RL Nature 390:249-256(1997).

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

CC EMBL; D84432; BAA12635.1; --

CC EMBL; Z99116; CAB14297.1; --

CC PIR; E69966; E69966

CC Subtilist; BGL1758; yqkC.

KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 79 AA; 9287 MW; 34852BDAAE4805D2 CRC64;

Query Match 27.1%; Score 52; DB 1; Length 79;
Best Local Similarity 28.1%; Pred. NO. 1.7;
Matches 9; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

QY 2 RSYRYKLLNWAYQQVQONKEDAWLEHVDVWRME 33
Db 23 QNHPTLLHWSIGGAESIKKDVLLQDEMTEF 54

RESULT 7

UN13_CAEEL

ID UN13_CAEEL STANDARD; PRT; 1813 AA.

AC P27715; Q17665; Q23512; Q81095;

DT 01-AUG-1992 (Rel. 23, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Phorbol ester/diacylglycerol-binding protein unc-13 (Uncoordinated

DE protein 13).

GN UNC-13 OR ZK524.2.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 2).

RC STRAIN=Bristol N2;

RX MEDLINE=91288538; PubMed=2062851;

RA Maruyama I.N., Brenner S.;

RT "A phorbol ester/diacylglycerol-binding protein encoded by the unc-13

RT gene of Caenorhabditis elegans.";

RL Proc. Natl. Acad. Sci. U.S.A. 88:5729-5733(1991).

RN [2]

RP PARTIAL SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).

RC STRAIN=Bristol N2;

RX MEDLINE=20483775; PubMed=11029047;

RA Kohn R.E., Duerr J.S., McManus J.R., Duke A., Rakow T.L., Maruyama H.,

RA Moulder G., Maruyama I.N., Barstead R.J., Rand J.B.;

RT "Expression of multiple UNC-13 proteins in the Caenorhabditis elegans

RT nervous system.";

RL Mol. Biol. Cell 11:3441-3452(2000).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Gardner A.E., Lloyd C.R.;

RT Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.

RN [4]

RP REVISIONS, AND ALTERNATIVE SPLICING.

RA Durbin R.;

RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.

RN [5]

RP ZINC AND PHORBOL-ESTERS BINDING.

RC MEDLINE=93075060; PubMed=1445255;

RA Ahmed S., Maruyama I.N., Kozma R., Lee J., Brenner S., Lim L.;

RT "The Caenorhabditis elegans unc-13 gene product is a phospholipid-

RT dependent high-affinity phorbol ester receptor.";

RL Biochem. J. 287:995-999(1992).

CC

CC -!- FUNCTION: May form part of a signal transduction pathway,

CC transducing the signal from diacylglycerol to effector functions.

CC One such function could be the release of neurotransmitter from

CC neurons.

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=5;

CC Name=1; Synonyms=ZK524.2a;

CC IsoId=P27715-1; Sequence=Displayed;

CC Name=2;

CC IsoId=P27715-2; Sequence=VSP_004484;

CC Name=3;

CC IsoId=P27715-3; Sequence=VSP_004485;

CC Name=4; Synonyms=ZK524.2b;

CC IsoId=P27715-4; Sequence=VSP_004486, VSP_004487;


```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D90913; BAA18354.1; -.
CC DR PIR; S75895; S75895.
CC DR HSSP; P08859; IGLC.
CC DR HAWAP; MF_00186; -.
CC DR InterPro; IPR000577; FGGY_kin.
CC DR InterPro; IPR005999; Glycerol_kin.
CC DR Pfam; PF00370; FGGY; 1.
CC DR Pfam; PF02782; FGGY_C; 1.
CC DR TIGRfams; TIGR01311; Glycerol_kin; 1.
CC DR PROSITE; PS00933; FGGY_KINASES_1; FALSE_NEG.
CC DR PROSITE; PS00445; FGGY_KINASES_2; 1.
CC KW Glycerol metabolism; Transferase; Kinase; ATP-binding;
CC Complete proteome.
CC FT NP_BIND 156 168 ATP (PROBABLE)
CC SQ SEQUENCE 495 AA; 54351 MW; 71AAC71E76816736 CRC64;

Query Match 26.6%; Score 51; DB 1; Length 495;
Best Local Similarity 34.4%; Pred. No. 18;
Matches 11; Conservative 7; Mismatches 10; Indels 4; Gaps 2;

Qy 4 YRYKLLNWAYQVOONKEDA-WIEHD---VWR 31
Db |::|::| | | | | | | | | | | | | |
26 YEGNIGVQAYKELTFYFKAGWVEHDALEINR 57

RESULT 9
ACAD ARATH STANDARD; PRT; 1017 AA.
ID ACAD ARATH
AC Q9LIK7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Potential calcium-transporting ATPase 13, plasma membrane-type
DE (EC 3.6.3.8) (Ca(2+)-ATPase isoform 13).
DE ACAL3 OR ATG22910 OR F5N5.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA MEDLINE=20363099; PubMed=10907853;
RA Kaneo T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC
RT and BAC clones."
RL DNA Res. 7:217-221 (2000).
CC -1- FUNCTION: This magnesium-dependent enzyme catalyzes the hydrolysis
CC of ATP coupled with the translocation of calcium from the cytosol
CC out of the cell or into organelles (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + Ca(2+) (Cis) = ADP + phosphate +
CC Ca(2+) (trans).
CC -1- ENZYME REGULATION: Activated by calmodulin (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- DOMAIN: The N-terminus contains an autoinhibitory calmodulin-
CC binding domain, which binds calmodulin in a calcium-dependent
CC fashion (By similarity).
CC -1- SIMILARITY: Belongs to the cation transport ATPases family (P-type
CC ATPases). Subfamily IIB.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
DR EMBL; AP001300; BAB03036.1; -.
DR HSSP; P04191; 1EUL.
DR InterPro; IPR006408; ATPase-IIB_Ca.
DR InterPro; IPR001757; ATPase_EI-E2.
DR InterPro; IPR006068; Cation_ATPase_C.
DR InterPro; IPR004014; Cation_ATPase_N.
DR InterPro; IPR008250; EI-E2_ATPase_Ieg.
DR InterPro; IPR005834; Hydrolase.
DR Pfam; PF00689; Cation_ATPase_C; 1.
DR Pfam; PF00690; Cation_ATPase_N; 1.
DR Pfam; PF00122; EI-E2_ATPase; 1.
DR Pfam; PF00702; Hydrolase; 1.
DR PRINTS; PR00119; CATATPASE.
DR TIGRfams; TIGR01517; ATPase-IIB_Ca; 1.
DR TIGRfams; TIGR01494; ATPase_P-type; 5.
DR PROSITE; PS00154; ATPase_EI-E2; 1.
KW Hydrolase; Calcium transport; Transmembrane; Phosphorylation;
KW ATP-binding; Metal-binding; Magnesium; Calmodulin-binding;
KW Multigene family; Hypothetical protein.
FT DOMAIN 1 147 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 148 168 POTENTIAL.
FT DOMAIN 169 186 LUMENAL (POTENTIAL).
FT TRANSMEM 187 207 POTENTIAL.
FT DOMAIN 208 336 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 337 356 POTENTIAL.
FT DOMAIN 357 393 LUMENAL (POTENTIAL).
FT TRANSMEM 394 411 POTENTIAL.
FT DOMAIN 412 802 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 803 821 POTENTIAL.
FT DOMAIN 822 832 LUMENAL (POTENTIAL).
FT TRANSMEM 833 853 POTENTIAL.
FT DOMAIN 854 873 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 874 896 POTENTIAL.
FT DOMAIN 897 905 LUMENAL (POTENTIAL).
FT TRANSMEM 906 926 POTENTIAL.
FT DOMAIN 927 944 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 945 966 POTENTIAL.
FT DOMAIN 967 976 LUMENAL (POTENTIAL).
FT TRANSMEM 977 998 POTENTIAL.
FT DOMAIN 999 1002 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1003 1020 CALMODULIN-BINDING (PROBABLE).
FT MOD_RES 20 31 PHOSPHORYLATION (BY SIMILARITY).
FT METAL 449 449 MAGNESIUM (BY SIMILARITY).
FT METAL 747 747 MAGNESIUM (BY SIMILARITY).
FT METAL 751 751 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 1017 AA; 112527 MW; 2F0265CCE8862916 CRC64;

Query Match 26.6%; Score 51; DB 1; Length 1017;
Best Local Similarity 34.6%; Pred. No. 40;
Matches 9; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

Qy 8 LLNWAYQVOONKEDAWIEHDVWRME 33
Db |::|::| | | | | | | | | | | | | |
516 ILNWAYEELNMGKVEIEHDVVRME 541

RESULT 10
TRUB PROMA STANDARD; PRT; 313 AA.
ID TRUB PROMA
AC P59850;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE tRNA pseudouridine synthase B (EC 4.2.1.70) (tRNA pseudouridine 55
DE synthase) (Pse55 synthase) (Pseudouridylate synthase) (Uracil
DE hydrolase)
DE TRUB OR PRO1428.
OS Prochlorococcus marinus.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1219;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SARG / COMP 1375 / SS120;
```

```

RX MEDLINE=22810154; PubMed=12917486;
RA Dufresne A., Salanoubat M., Partensky F., Artiguenave F., Axmann I.M.,
RA Barbe V., Duprat S., Galperin M.Y., Koonin E.V., Le Gall F.,
RA Makarova K.S., Ostrowski M., Oztas S., Robert C., Rogozin I.B.,
RA Scanlan D.J., Tandeau de Marsac N., Weissenbach J., Wincker P.,
RA Wolf Y.I., Hess W.R.,
RT "Genome sequence of the cyanobacterium Prochlorococcus marinus SS120,
RT a nearly minimal oxic phototrophic genome."
RL Proc. Natl. Acad. Sci. U.S.A. 100:10020-10025(2003).
CC -!- FUNCTION: Responsible for synthesis of pseudouridine from
CC uracil-55 in the psi GC loop of transfer RNAs (By similarity).
CC -!- CATALYTIC ACTIVITY: Uracil + D-ribose 5-phosphate = pseudouridine
CC 5'-phosphate + H(2)O.
CC -!- SIMILARITY: Belongs to the pseudouridine synthase trnB family.
CC Subfamily 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB017165; AAQ00472.1; -.
DR HAMAP; MF_01080; -.
KW tRNA processing; Lyase; Complete proteome.
FT ACT SITE 42 42 BY SIMILARITY.
SQ SEQUENCE 313 AA; 35024 MW; 34BF5DEB05FF172B CRC64;

Query Match 26.3%; Score 50.5; DB 1; Length 313;
Best Local Similarity 42.3%; Pred. No. 13;
Matches 11; Conservative 4; Mismatches 4; Indels 7; Gaps 1;

QY 4 YRYKILNWAYQQVQONKEDAWIEHDV 29
DB 150 YRIKILN-----NKKDGTIDLEV 168

RESULT 11
ALP4_SCHPO STANDARD; PRT; 784 AA.
AC Q9Y705;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Spindle pole body component alp4 (Altered polarity protein 4).
GN ALP4 OR SPBC365.15.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]_
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=972;
RX MEDLINE=20532503; PubMed=11080156;
RA Wady L., Toda T.,
RT "The fission yeast gamma-tubulin complex is required in G(1) phase and
RT is a component of the spindle assembly checkpoint."
RL ENBO J. 19:6098-6111(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,

```

```

RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Timmons R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovsky G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -!- FUNCTION: Component of the gamma tubule complex that is required
CC for the regulation of both interphase microtubules and mitotic
CC bipolar spindles.
CC -!- SUBCELLULAR LOCATION: Spindle pole body and the microtubule
CC organizing center (MTOC).
CC -!- SIMILARITY: Belongs to the GCP family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL078627; CAB44767.1; -.
DR EMBL; AB026664; BAA77269.1; -.
DR PIR; T43510; T43510.
DR GeneDB; SPombe: SPBC365.15; -.
DR InterPro; IPR007259; Spc97_Spc98.
DR Pfam; PF04130; Spc97_Spc98; 1.
KW Microtubule; Mitosis.
SQ SEQUENCE 784 AA; 90157 MW; E529CE217FFDA2B2 CRC64;

Query Match 26.0%; Score 50; DB 1; Length 784;
Best Local Similarity 31.2%; Pred. No. 41;
Matches 10; Conservative 8; Mismatches 8; Indels 6; Gaps 2;

QY 4 YRYKILNWAYQQVQONKEDAWIEH---DVWRM 32
DB 592 FRYFLL---LRHVMQLENSVQHSKNSARLL 620

RESULT 12
SVJ1_RAT STANDARD; PRT; 1574 AA.
ID SVJ1_RAT
AC Q629I0; O89092; Q629I1; Q81028;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Synaptotagmin 1 (EC 3.1.3.36) (Synaptic inositol-1,4,5-trisphosphate 5-
DE phosphatase 1).
GN SVNJ1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]_
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC STRAIN=Sprague-Dawley; Tissue=Brain;
RX MEDLINE=96149250; PubMed=8552192;
RA McPherson P.S., Garcia E.P., Slepnev V.I., David C., Zhang X.,
RA Grabs D., Sossin W.S., Bauerfeind R., Nemoto Y., De Camilli P.;
RT "A presynaptic inositol-5-phosphatase."
RT Nature 379:353-357(1996).

```



```

RN RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=20001959; PubMed=10529403;
RA Okuda T., Matsuda S., Nakatsugawa S., Ichigotani Y., Iwahashi N.,
RT Takahashi M., Ishigaki T., Hamaguchi M.;
RT "Molecular cloning of macrophin, a human homologue of Drosophila
RT kakapo with a close structural similarity to plectin and dystrophin.";
RJ Biochem. Biophys. Res. Commun. 264:568-574 (1999).
RN [2]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=20026884; PubMed=10559237;
RA Sun Y., Zhang J., Kraeft S.-K., Auclair D., Chang M.-S., Liu Y.,
RA Sutherland R., Sargia R., Griffin J.D., Ferland L.H., Chen L.B.;
RT "Molecular cloning and characterization of human trabeculin-alpha, a
RT grant protein defining a new family of actin-binding proteins.";
RJ J. Biol. Chem. 274:33522-33530 (1999).
RN [3]
RN SEQUENCE OF 182-4812 FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=2183812; PubMed=11845288;
RA Gong T.-W.L., Besirli C.G., Lomax M.I.;
RT "WACFI gene structure: a hybrid of plectin and dystrophin.";
RJ Mamm. Genome 12:852-861 (2001).
RN [4]
RN SEQUENCE OF 868-2350 FROM N.A.
RX TISSUE=Brain;
RA Nagase T., Ishikawa K.-I., Kikuno R., Hirose M., Nomura N.,
RA Chara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RJ DNA Res. 6:337-345 (1999).
RN [5]
RN SEQUENCE OF 1544-5057 FROM N.A.
RA Corby N.;
RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RN SEQUENCE OF 3312-5430 FROM N.A.
RX TISSUE=Brain;
RA MEDLINE=22158633; PubMed=12168954;
RA Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;
RT "Construction of expression-ready cDNA clones for KIAA genes: manual
RT curation of 330 KIAA cDNA clones.";
RJ DNA Res. 9:99-106 (2002).
RN [7]
RN SEQUENCE OF 3734-5430 FROM N.A.
RX TISSUE=Brain;
RA MEDLINE=98116662; PubMed=9455484;
RA Seki N., Ohira M., Nagase T., Ishikawa K.-I., Miyajima N.,
RA Nakajima D., Nomura N., Ohara O.;
RT "Characterization of cDNA clones in size-fractionated cDNA libraries
RT from human brain.";
RJ DNA Res. 4:345-349 (1997).
RN [8]
RN FUNCTION: F-actin-binding protein which may play a role in cross-
RN linking actin to other cytoskeletal proteins. Also binds to
RN microtubules (By similarity).
RN
RN SUBCELLULAR LOCATION: Cytoplasmic.
RN
RN ALTERNATIVE PRODUCTS:
RN Name=2;
RN IsoId=Q9UPN3-2; Sequence=Displayed;
RN Name=1;
RN IsoId=Q9UPN3-1; Sequence=VSP_007341;
RN Name=3;
RN IsoId=Q9UPN3-3; Sequence=Not described;
RN Name=4;
RN IsoId=Q96PK2-1; Sequence=External;
RN
RN TISSUE SPECIFICITY: Ubiquitously expressed.
RN
RN SIMILARITY: Belongs to the plectin or cytolinker family.
RN
RN SIMILARITY: Contains 1 actin-binding domain.
RN
RN SIMILARITY: Contains 2 calponin-homology (CH) domains.
RN
RN SIMILARITY: Contains 2 EF-hand calcium-binding domains.
RN
RN SIMILARITY: Contains 1 SH3 domain.
RN
RN SIMILARITY: Contains 37 spectrin repeats.
RN
RN SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (see http://www.isb-sib.ch/announce/
RN or send an email to license@isb-sib.ch).
RN
RN EMBL; AB029290; BAA83821.1; -
RN EMBL; AF141968; AAF06360.1; -
RN EMBL; AF325341; AAL39000.1; -
RN EMBL; AF325330; AAL39000.1; JOINED.
RN EMBL; AF325331; AAL39000.1; JOINED.
RN EMBL; AF325332; AAL39000.1; JOINED.
RN EMBL; AF325333; AAL39000.1; JOINED.
RN EMBL; AF325334; AAL39000.1; JOINED.
RN EMBL; AF325335; AAL39000.1; JOINED.
RN EMBL; AF325336; AAL39000.1; JOINED.
RN EMBL; AF325337; AAL39000.1; JOINED.
RN EMBL; AF325340; AAL39000.1; JOINED.
RN EMBL; AB033077; BAA86565.1; -
RN EMBL; AL137853; CAC15920.1; -
RN EMBL; AB007934; BAA32310.2; -
RN PIR; T00079; T00079.
RN HSP; Q01082; IBKR.
RN Genew; HGNC:13664; MACF1.
RN GO; GO:0005856; Cytoskeleton; NAS.
RN GO; GO:0003780; F-actin cross-linking activity; NAS.
RN GO; GO:0005509; F-actin ion binding; NAS.
RN GO; GO:0008017; F-actin microtubule binding; NAS.
RN InterPro; IPR001589; Actin-like actin.
RN InterPro; IPR001715; Calponin-like.
RN InterPro; IPR002048; EF-hand.
RN InterPro; IPR003108; GAS2.
RN InterPro; IPR001452; SH3.
RN InterPro; IPR002017; Spectrin.
RN Pfam; PF00307; CH; 2.
RN Pfam; PF00036; ehand; 2.
RN Pfam; PF02187; GAS2; 1.
RN Pfam; PF00435; spectrin; 27.
RN ProDom; PD000012; EF-hand; 1.
RN SMART; SM00033; CH; 2.
RN SMART; SM00054; EFR; 2.
RN SMART; SM00243; GAS2; 1.
RN SMART; SM0150; SPEC; 36.
RN PROSITE; PS00019; ACTININ_1; 1.
RN PROSITE; PS00020; ACTININ_2; FALSE_NEG.
RN PROSITE; PS00021; CH; 2.
RN PROSITE; PS00018; EF_HAND; 2.
RN PROSITE; PS00002; SH3; FALSE_NEG.
RN Actin-binding; Cytoskeleton; Calcium; Calcium-binding; Repeat;
RN SH3 domain; Coiled coil; Alternative splicing.
RN DOMAIN 1 295 ACTIN-BINDING.
RN DOMAIN 78 181 CH 1.
RN DOMAIN 194 295 CH 2.
RN DOMAIN 243 265 COILED COIL (POTENTIAL).
RN DOMAIN 477 529 COILED COIL (POTENTIAL).
RN DOMAIN 722 751 COILED COIL (POTENTIAL).
RN DOMAIN 816 843 COILED COIL (POTENTIAL).
RN DOMAIN 1013 1118 COILED COIL (POTENTIAL).
RN DOMAIN 1164 1191 COILED COIL (POTENTIAL).
RN DOMAIN 1399 1690 COILED COIL (POTENTIAL).
RN DOMAIN 1780 1843 COILED COIL (POTENTIAL).
RN DOMAIN 1975 2005 COILED COIL (POTENTIAL).
RN DOMAIN 2039 2312 COILED COIL (POTENTIAL).
RN DOMAIN 2385 2417 COILED COIL (POTENTIAL).
RN DOMAIN 2544 2695 COILED COIL (POTENTIAL).
RN DOMAIN 2760 2838 COILED COIL (POTENTIAL).
RN DOMAIN 2911 3001 COILED COIL (POTENTIAL).
RN DOMAIN 3130 3164 COILED COIL (POTENTIAL).
RN DOMAIN 3244 3277 COILED COIL (POTENTIAL).
```

```
FT DOMAIN 3418 3482 COILED COIL (POTENTIAL) .
FT DOMAIN 3596 3666 COILED COIL (POTENTIAL) .
FT DOMAIN 3786 3806 COILED COIL (POTENTIAL) .
FT DOMAIN 3852 3931 COILED COIL (POTENTIAL) .
FT DOMAIN 3967 3997 COILED COIL (POTENTIAL) .
FT DOMAIN 4084 4218 COILED COIL (POTENTIAL) .
FT DOMAIN 4343 4378 COILED COIL (POTENTIAL) .
FT DOMAIN 4408 4437 COILED COIL (POTENTIAL) .
FT DOMAIN 4468 4498 COILED COIL (POTENTIAL) .
FT DOMAIN 4907 4935 COILED COIL (POTENTIAL) .
FT DOMAIN 5044 5067 COILED COIL (POTENTIAL) .
FT REPEAT 314 355 SPECTRIN 1.
FT REPEAT 591 623 SPECTRIN 2.
FT REPEAT 680 784 SPECTRIN 3.
FT REPEAT 786 800 SPECTRIN 4.
FT DOMAIN 871 923 SH3.
FT REPEAT 1250 1272 SPECTRIN 5.
FT REPEAT 1287 1342 SPECTRIN 6.
FT REPEAT 1455 1534 SPECTRIN 7.
FT REPEAT 1547 1659 SPECTRIN 8.
FT REPEAT 1815 1891 SPECTRIN 9.
FT REPEAT 1932 2042 SPECTRIN 10.
FT REPEAT 2260 2280 SPECTRIN 11.
FT REPEAT 2372 2395 SPECTRIN 12.
FT REPEAT 2398 2507 SPECTRIN 13.
FT REPEAT 2510 2618 SPECTRIN 14.
FT REPEAT 2621 2728 SPECTRIN 15.
FT REPEAT 2731 2838 SPECTRIN 16.
FT REPEAT 2841 2945 SPECTRIN 17.
FT REPEAT 2987 3024 SPECTRIN 18.
FT REPEAT 3136 3163 SPECTRIN 19.
FT REPEAT 3187 3274 SPECTRIN 20.
FT REPEAT 3277 3383 SPECTRIN 21.
FT REPEAT 3386 3492 SPECTRIN 22.
FT REPEAT 3495 3601 SPECTRIN 23.
FT REPEAT 3604 3673 SPECTRIN 24.
FT REPEAT 3713 3819 SPECTRIN 25.
FT REPEAT 3832 3927 SPECTRIN 26.
FT REPEAT 3982 4043 SPECTRIN 27.
FT REPEAT 4046 4152 SPECTRIN 28.
FT REPEAT 4155 4262 SPECTRIN 29.

Query Match 26.0%; Score 50; DB 1; Length 5430;
Best Local Similarity 40.7%; Pred. No. 3.6e+02;
Matches 11; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 5 RYKLNWAYQQVQNKEDAWIEHDVWR 31
Db 5048 RQKLNDAIDRLLEELKEFANFDVWR 5074

RESULT 14
MAC4_HUMAN
ID MAC4_HUMAN STANDARD; PRT; 5938 AA.
AC Q96PK2; Q8WXY1;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Microtubule-actin crosslinking factor 1, isoform 4.
GN MACF1 OR ACF7 OR ABP620 OR KIAA0465 OR KIAA1251.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RX MEDLINE=21833812; PubMed=11845288;
RA Gong T.-W.L., Besirli C.G., Lomax M.I.;
RT "MACF1 gene structure: a hybrid of plectin and dystrophin.";
RL Mamm. Genome 12:852-861(2001).
CC -!- FUNCTION: May play a role in cross-linking cytoskeletal proteins
CC by binding intermediate filaments to the N-terminal plectin
CC repeats and microtubules to the C-terminus.
CC
```

```
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity) .
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=4;
CC IsoId=Q96PK2-1; Sequence=Displayed;
CC Name=1;
CC IsoId=Q9UPN3-1; Sequence=External;
CC Name=2;
CC IsoId=Q9UPN3-2; Sequence=External;
CC Name=3;
CC IsoId=Q9UPN3-3; Sequence=External;
CC -!- TISSUE SPECIFICITY: Expressed in heart, lung, pituitary and
CC placenta, not found in brain, kidney, liver, pancreas or skeletal
CC muscle.
CC -!- SIMILARITY: Belongs to the plectin or cytolinker family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC -!- SIMILARITY: Contains 19 plectrin repeats.
CC -!- SIMILARITY: Contains 32 spectrin repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF317696; AAL09459.1; -.
CC EMBL; AF325341; AAL38997.1; -.
CC EMBL; AF325333; AAL38997.1; JOINED.
CC EMBL; AF325334; AAL38997.1; JOINED.
CC EMBL; AF325335; AAL38997.1; JOINED.
CC EMBL; AF325336; AAL38997.1; JOINED.
CC EMBL; AF325339; AAL38997.1; JOINED.
CC EMBL; AF325340; AAL38997.1; JOINED.
CC DR GO; GO:0005856; C:cytoskeleton; ISS.
CC DR GO; GO:0005509; F:calcium ion binding; ISS.
CC DR GO; GO:0008017; F:microtubule binding; ISS.
CC DR InterPro; IPR002048; EF-hand.
CC DR InterPro; IPR003108; GAS2.
CC DR InterPro; IPR001101; Plectrin_repeat.
CC DR InterPro; IPR002017; Spectrin.
CC Pfam; PF00036; ehand; 2.
CC Pfam; PF02187; GAS2; 1.
CC Pfam; PF00681; Plectrin; 11.
CC Pfam; PF00435; spectrin; 26.
CC ProDom; PDC00012; EF-hand; 1.
CC DR SMART; SM00054; EFh; 2.
CC DR SMART; SM00243; GAS2; 1.
CC DR SMART; SM00250; PLEC; 19.
CC DR SMART; SM00150; SPEC; 32.
CC PROSITE; PS00018; EF_HAND; 2.
KW Cytoskeleton; Calcium; Calcium-binding; Repeat; Coiled coil;
KW Alternative splicing.
FT DOMAIN 1830 1936 COILED COIL (POTENTIAL) .
FT DOMAIN 2001 2192 COILED COIL (POTENTIAL) .
FT DOMAIN 2282 2345 COILED COIL (POTENTIAL) .
FT DOMAIN 2477 2507 COILED COIL (POTENTIAL) .
FT DOMAIN 2541 2654 COILED COIL (POTENTIAL) .
FT DOMAIN 2686 2814 COILED COIL (POTENTIAL) .
FT DOMAIN 2887 2919 COILED COIL (POTENTIAL) .
FT DOMAIN 3046 3197 COILED COIL (POTENTIAL) .
FT DOMAIN 3262 3503 COILED COIL (POTENTIAL) .
FT DOMAIN 3632 3666 COILED COIL (POTENTIAL) .
FT DOMAIN 3746 3779 COILED COIL (POTENTIAL) .
FT DOMAIN 3920 3984 COILED COIL (POTENTIAL) .
FT DOMAIN 4098 4168 COILED COIL (POTENTIAL) .
FT DOMAIN 4288 4308 COILED COIL (POTENTIAL) .
FT DOMAIN 4354 4386 COILED COIL (POTENTIAL) .
FT DOMAIN 4397 4433 COILED COIL (POTENTIAL) .
FT DOMAIN 4469 4489 COILED COIL (POTENTIAL) .
FT DOMAIN 4586 4720 COILED COIL (POTENTIAL) .
FT DOMAIN 4845 4880 COILED COIL (POTENTIAL) .
```

```
FT DOMAIN 4910 4939 COILED COIL (POTENTIAL) .
FT DOMAIN 4970 5000 COILED COIL (POTENTIAL) .
FT DOMAIN 5409 5437 COILED COIL (POTENTIAL) .
FT DOMAIN 5546 5569 COILED COIL (POTENTIAL) .
FT CA_BIND 5598 5610 EF-HAND 1 (POTENTIAL) .
FT CA_BIND 5634 5646 EF-HAND 2 (POTENTIAL) .
FT REPEAT 12 49 PLECTIN 1.
FT REPEAT 53 88 PLECTIN 2.
FT REPEAT 89 126 PLECTIN 3.
FT REPEAT 130 164 PLECTIN 4.
FT REPEAT 166 202 PLECTIN 5.
FT REPEAT 203 240 PLECTIN 6.
FT REPEAT 243 278 PLECTIN 7.
FT REPEAT 279 316 PLECTIN 8.
FT REPEAT 318 354 PLECTIN 9.
FT REPEAT 354 388 PLECTIN 10.
FT REPEAT 388 412 PLECTIN 11.
FT REPEAT 412 436 PLECTIN 12.
FT REPEAT 436 460 PLECTIN 13.
FT REPEAT 460 484 PLECTIN 14.
FT REPEAT 484 508 PLECTIN 15.
FT REPEAT 508 532 PLECTIN 16.
FT REPEAT 532 556 PLECTIN 17.
FT REPEAT 556 580 PLECTIN 18.
FT REPEAT 580 604 PLECTIN 19.
FT REPEAT 604 628 SPECTRIN 1.
FT REPEAT 628 652 SPECTRIN 2.
FT REPEAT 652 676 SPECTRIN 3.
FT REPEAT 676 700 SPECTRIN 4.
FT REPEAT 700 724 SPECTRIN 5.
FT REPEAT 724 748 SPECTRIN 6.
FT REPEAT 748 772 SPECTRIN 7.
FT REPEAT 772 796 SPECTRIN 8.
FT REPEAT 796 820 SPECTRIN 9.
FT REPEAT 820 844 SPECTRIN 10.
FT REPEAT 844 868 SPECTRIN 11.
FT REPEAT 868 892 SPECTRIN 12.
FT REPEAT 892 916 SPECTRIN 13.
FT REPEAT 916 940 SPECTRIN 14.
FT REPEAT 940 964 SPECTRIN 15.
FT REPEAT 964 988 SPECTRIN 16.
FT REPEAT 988 1012 SPECTRIN 17.
FT REPEAT 1012 1036 SPECTRIN 18.
FT REPEAT 1036 1060 SPECTRIN 19.
FT REPEAT 1060 1084 SPECTRIN 20.
FT REPEAT 1084 1108 SPECTRIN 21.
FT REPEAT 1108 1132 SPECTRIN 22.
FT REPEAT 1132 1156 SPECTRIN 23.
FT REPEAT 1156 1180 SPECTRIN 24.
FT REPEAT 1180 1204 SPECTRIN 25.
FT REPEAT 1204 1228 SPECTRIN 26.
FT REPEAT 1228 1252 SPECTRIN 27.
FT REPEAT 1252 1276 SPECTRIN 28.
FT REPEAT 1276 1300 SPECTRIN 29.
FT REPEAT 1300 1324 SPECTRIN 30.
FT REPEAT 1324 1348 SPECTRIN 31.
FT REPEAT 1348 1372 SPECTRIN 32.
FT REPEAT 1372 1396 SPECTRIN 33.
FT REPEAT 1396 1420 SPECTRIN 34.
FT REPEAT 1420 1444 SPECTRIN 35.
FT REPEAT 1444 1468 SPECTRIN 36.
FT REPEAT 1468 1492 SPECTRIN 37.
FT REPEAT 1492 1516 SPECTRIN 38.
FT REPEAT 1516 1540 SPECTRIN 39.
FT REPEAT 1540 1564 SPECTRIN 40.
FT REPEAT 1564 1588 SPECTRIN 41.
FT REPEAT 1588 1612 SPECTRIN 42.
FT REPEAT 1612 1636 SPECTRIN 43.
FT REPEAT 1636 1660 SPECTRIN 44.
FT REPEAT 1660 1684 SPECTRIN 45.
FT REPEAT 1684 1708 SPECTRIN 46.
FT REPEAT 1708 1732 SPECTRIN 47.
FT REPEAT 1732 1756 SPECTRIN 48.
FT REPEAT 1756 1780 SPECTRIN 49.
FT REPEAT 1780 1804 SPECTRIN 50.
FT REPEAT 1804 1828 SPECTRIN 51.
FT REPEAT 1828 1852 SPECTRIN 52.
FT REPEAT 1852 1876 SPECTRIN 53.
FT REPEAT 1876 1900 SPECTRIN 54.
FT REPEAT 1900 1924 SPECTRIN 55.
FT REPEAT 1924 1948 SPECTRIN 56.
FT REPEAT 1948 1972 SPECTRIN 57.
FT REPEAT 1972 1996 SPECTRIN 58.
FT REPEAT 1996 2020 SPECTRIN 59.
FT REPEAT 2020 2044 SPECTRIN 60.
FT REPEAT 2044 2068 SPECTRIN 61.
FT REPEAT 2068 2092 SPECTRIN 62.
FT REPEAT 2092 2116 SPECTRIN 63.
FT REPEAT 2116 2140 SPECTRIN 64.
FT REPEAT 2140 2164 SPECTRIN 65.
FT REPEAT 2164 2188 SPECTRIN 66.
FT REPEAT 2188 2212 SPECTRIN 67.
FT REPEAT 2212 2236 SPECTRIN 68.
FT REPEAT 2236 2260 SPECTRIN 69.
FT REPEAT 2260 2284 SPECTRIN 70.
FT REPEAT 2284 2308 SPECTRIN 71.
FT REPEAT 2308 2332 SPECTRIN 72.
FT REPEAT 2332 2356 SPECTRIN 73.
FT REPEAT 2356 2380 SPECTRIN 74.
FT REPEAT 2380 2404 SPECTRIN 75.
FT REPEAT 2404 2428 SPECTRIN 76.
FT REPEAT 2428 2452 SPECTRIN 77.
FT REPEAT 2452 2476 SPECTRIN 78.
FT REPEAT 2476 2500 SPECTRIN 79.
FT REPEAT 2500 2524 SPECTRIN 80.
FT REPEAT 2524 2548 SPECTRIN 81.
FT REPEAT 2548 2572 SPECTRIN 82.
FT REPEAT 2572 2596 SPECTRIN 83.
FT REPEAT 2596 2620 SPECTRIN 84.
FT REPEAT 2620 2644 SPECTRIN 85.
FT REPEAT 2644 2668 SPECTRIN 86.
FT REPEAT 2668 2692 SPECTRIN 87.
FT REPEAT 2692 2716 SPECTRIN 88.
FT REPEAT 2716 2740 SPECTRIN 89.
FT REPEAT 2740 2764 SPECTRIN 90.
FT REPEAT 2764 2788 SPECTRIN 91.
FT REPEAT 2788 2812 SPECTRIN 92.
FT REPEAT 2812 2836 SPECTRIN 93.
FT REPEAT 2836 2860 SPECTRIN 94.
FT REPEAT 2860 2884 SPECTRIN 95.
FT REPEAT 2884 2908 SPECTRIN 96.
FT REPEAT 2908 2932 SPECTRIN 97.
FT REPEAT 2932 2956 SPECTRIN 98.
FT REPEAT 2956 2980 SPECTRIN 99.
FT REPEAT 2980 3004 SPECTRIN 100.
FT REPEAT 3004 3028 SPECTRIN 101.
FT REPEAT 3028 3052 SPECTRIN 102.
FT REPEAT 3052 3076 SPECTRIN 103.
FT REPEAT 3076 3100 SPECTRIN 104.
FT REPEAT 3100 3124 SPECTRIN 105.
FT REPEAT 3124 3148 SPECTRIN 106.
FT REPEAT 3148 3172 SPECTRIN 107.
FT REPEAT 3172 3196 SPECTRIN 108.
FT REPEAT 3196 3220 SPECTRIN 109.
FT REPEAT 3220 3244 SPECTRIN 110.
FT REPEAT 3244 3268 SPECTRIN 111.
FT REPEAT 3268 3292 SPECTRIN 112.
FT REPEAT 3292 3316 SPECTRIN 113.
FT REPEAT 3316 3340 SPECTRIN 114.
FT REPEAT 3340 3364 SPECTRIN 115.
FT REPEAT 3364 3388 SPECTRIN 116.
FT REPEAT 3388 3412 SPECTRIN 117.
FT REPEAT 3412 3436 SPECTRIN 118.
FT REPEAT 3436 3460 SPECTRIN 119.
FT REPEAT 3460 3484 SPECTRIN 120.
FT REPEAT 3484 3508 SPECTRIN 121.
FT REPEAT 3508 3532 SPECTRIN 122.
FT REPEAT 3532 3556 SPECTRIN 123.
FT REPEAT 3556 3580 SPECTRIN 124.
FT REPEAT 3580 3604 SPECTRIN 125.
FT REPEAT 3604 3628 SPECTRIN 126.
FT REPEAT 3628 3652 SPECTRIN 127.
FT REPEAT 3652 3676 SPECTRIN 128.
FT REPEAT 3676 3700 SPECTRIN 129.
FT REPEAT 3700 3724 SPECTRIN 130.
FT REPEAT 3724 3748 SPECTRIN 131.
FT REPEAT 3748 3772 SPECTRIN 132.
FT REPEAT 3772 3796 SPECTRIN 133.
FT REPEAT 3796 3820 SPECTRIN 134.
FT REPEAT 3820 3844 SPECTRIN 135.
FT REPEAT 3844 3868 SPECTRIN 136.
FT REPEAT 3868 3892 SPECTRIN 137.
FT REPEAT 3892 3916 SPECTRIN 138.
FT REPEAT 3916 3940 SPECTRIN 139.
FT REPEAT 3940 3964 SPECTRIN 140.
FT REPEAT 3964 3988 SPECTRIN 141.
FT REPEAT 3988 4012 SPECTRIN 142.
FT REPEAT 4012 4036 SPECTRIN 143.
FT REPEAT 4036 4060 SPECTRIN 144.
FT REPEAT 4060 4084 SPECTRIN 145.
FT REPEAT 4084 4108 SPECTRIN 146.
FT REPEAT 4108 4132 SPECTRIN 147.
FT REPEAT 4132 4156 SPECTRIN 148.
FT REPEAT 4156 4180 SPECTRIN 149.
FT REPEAT 4180 4204 SPECTRIN 150.
FT REPEAT 4204 4228 SPECTRIN 151.
FT REPEAT 4228 4252 SPECTRIN 152.
FT REPEAT 4252 4276 SPECTRIN 153.
FT REPEAT 4276 4300 SPECTRIN 154.
FT REPEAT 4300 4324 SPECTRIN 155.
FT REPEAT 4324 4348 SPECTRIN 156.
FT REPEAT 4348 4372 SPECTRIN 157.
FT REPEAT 4372 4396 SPECTRIN 158.
FT REPEAT 4396 4420 SPECTRIN 159.
FT REPEAT 4420 4444 SPECTRIN 160.
FT REPEAT 4444 4468 SPECTRIN 161.
FT REPEAT 4468 4492 SPECTRIN 162.
FT REPEAT 4492 4516 SPECTRIN 163.
FT REPEAT 4516 4540 SPECTRIN 164.
FT REPEAT 4540 4564 SPECTRIN 165.
FT REPEAT 4564 4588 SPECTRIN 166.
FT REPEAT 4588 4612 SPECTRIN 167.
FT REPEAT 4612 4636 SPECTRIN 168.
FT REPEAT 4636 4660 SPECTRIN 169.
FT REPEAT 4660 4684 SPECTRIN 170.
FT REPEAT 4684 4708 SPECTRIN 171.
FT REPEAT 4708 4732 SPECTRIN 172.
FT REPEAT 4732 4756 SPECTRIN 173.
FT REPEAT 4756 4780 SPECTRIN 174.
FT REPEAT 4780 4804 SPECTRIN 175.
FT REPEAT 4804 4828 SPECTRIN 176.
FT REPEAT 4828 4852 SPECTRIN 177.
FT REPEAT 4852 4876 SPECTRIN 178.
FT REPEAT 4876 4900 SPECTRIN 179.
FT REPEAT 4900 4924 SPECTRIN 180.
FT REPEAT 4924 4948 SPECTRIN 181.
FT REPEAT 4948 4972 SPECTRIN 182.
FT REPEAT 4972 4996 SPECTRIN 183.
FT REPEAT 4996 5020 SPECTRIN 184.
FT REPEAT 5020 5044 SPECTRIN 185.
FT REPEAT 5044 5068 SPECTRIN 186.
FT REPEAT 5068 5092 SPECTRIN 187.
FT REPEAT 5092 5116 SPECTRIN 188.
FT REPEAT 5116 5140 SPECTRIN 189.
FT REPEAT 5140 5164 SPECTRIN 190.
FT REPEAT 5164 5188 SPECTRIN 191.
FT REPEAT 5188 5212 SPECTRIN 192.
FT REPEAT 5212 5236 SPECTRIN 193.
FT REPEAT 5236 5260 SPECTRIN 194.
FT REPEAT 5260 5284 SPECTRIN 195.
FT REPEAT 5284 5308 SPECTRIN 196.
FT REPEAT 5308 5332 SPECTRIN 197.
FT REPEAT 5332 5356 SPECTRIN 198.
FT REPEAT 5356 5380 SPECTRIN 199.
FT REPEAT 5380 5404 SPECTRIN 200.
FT REPEAT 5404 5428 SPECTRIN 201.
FT REPEAT 5428 5452 SPECTRIN 202.
FT REPEAT 5452 5476 SPECTRIN 203.
FT REPEAT 5476 5500 SPECTRIN 204.
FT REPEAT 5500 5524 SPECTRIN 205.
FT REPEAT 5524 5548 SPECTRIN 206.
FT REPEAT 5548 5572 SPECTRIN 207.
FT REPEAT 5572 5596 SPECTRIN 208.
FT REPEAT 5596 5620 SPECTRIN 209.
FT REPEAT 5620 5644 SPECTRIN 210.
FT REPEAT 5644 5668 SPECTRIN 211.
FT REPEAT 5668 5692 SPECTRIN 212.
FT REPEAT 5692 5716 SPECTRIN 213.
FT REPEAT 5716 5740 SPECTRIN 214.
FT REPEAT 5740 5764 SPECTRIN 215.
FT REPEAT 5764 5788 SPECTRIN 216.
FT REPEAT 5788 5812 SPECTRIN 217.
FT REPEAT 5812 5836 SPECTRIN 218.
FT REPEAT 5836 5860 SPECTRIN 219.
FT REPEAT 5860 5884 SPECTRIN 220.
FT REPEAT 5884 5908 SPECTRIN 221.
FT REPEAT 5908 5932 SPECTRIN 222.
FT REPEAT 5932 5956 SPECTRIN 223.
FT REPEAT 5956 5980 SPECTRIN 224.
FT REPEAT 5980 6004 SPECTRIN 225.
FT REPEAT 6004 6028 SPECTRIN 226.
FT REPEAT 6028 6052 SPECTRIN 227.
FT REPEAT 6052 6076 SPECTRIN 228.
FT REPEAT 6076 6100 SPECTRIN 229.
FT REPEAT 6100 6124 SPECTRIN 230.
FT REPEAT 6124 6148 SPECTRIN 231.
FT REPEAT 6148 6172 SPECTRIN 232.
FT REPEAT 6172 6196 SPECTRIN 233.
FT REPEAT 6196 6220 SPECTRIN 234.
FT REPEAT 6220 6244 SPECTRIN 235.
FT REPEAT 6244 6268 SPECTRIN 236.
FT REPEAT 6268 6292 SPECTRIN 237.
FT REPEAT 6292 6316 SPECTRIN 238.
FT REPEAT 6316 6340 SPECTRIN 239.
FT REPEAT 6340 6364 SPECTRIN 240.
FT REPEAT 6364 6388 SPECTRIN 241.
FT REPEAT 6388 6412 SPECTRIN 242.
FT REPEAT 6412 6436 SPECTRIN 243.
FT REPEAT 6436 6460 SPECTRIN 244.
FT REPEAT 6460 6484 SPECTRIN 245.
FT REPEAT 6484 6508 SPECTRIN 246.
FT REPEAT 6508 6532 SPECTRIN 247.
FT REPEAT 6532 6556 SPECTRIN 248.
FT REPEAT 6556 6580 SPECTRIN 249.
FT REPEAT 6580 6604 SPECTRIN 250.
FT REPEAT 6604 6628 SPECTRIN 251.
FT REPEAT 6628 6652 SPECTRIN 252.
FT REPEAT 6652 6676 SPECTRIN 253.
FT REPEAT 6676 6700 SPECTRIN 254.
FT REPEAT 6700 6724 SPECTRIN 255.
FT REPEAT 6724 6748 SPECTRIN 256.
FT REPEAT 6748 6772 SPECTRIN 257.
FT REPEAT 6772 6796 SPECTRIN 258.
FT REPEAT 6796 6820 SPECTRIN 259.
FT REPEAT 6820 6844 SPECTRIN 260.
FT REPEAT 6844 6868 SPECTRIN 261.
FT REPEAT 6868 6892 SPECTRIN 262.
FT REPEAT 6892 6916 SPECTRIN 263.
FT REPEAT 6916 6940 SPECTRIN 264.
FT REPEAT 6940 6964 SPECTRIN 265.
FT REPEAT 6964 6988 SPECTRIN 266.
FT REPEAT 6988 7012 SPECTRIN 267.
FT REPEAT 7012 7036 SPECTRIN 268.
FT REPEAT 7036 7060 SPECTRIN 269.
FT REPEAT 7060 7084 SPECTRIN 270.
FT REPEAT 7084 7108 SPECTRIN 271.
FT REPEAT 7108 7132 SPECTRIN 272.
FT REPEAT 7132 7156 SPECTRIN 273.
FT REPEAT 7156 7180 SPECTRIN 274.
FT REPEAT 7180 7204 SPECTRIN 275.
FT REPEAT 7204 7228 SPECTRIN 276.
FT REPEAT 7228 7252 SPECTRIN 277.
FT REPEAT 7252 7276 SPECTRIN 278.
FT REPEAT 7276 7300 SPECTRIN 279.
FT REPEAT 7300 7324 SPECTRIN 280.
FT REPEAT 7324 7348 SPECTRIN 281.
FT REPEAT 7348 7372 SPECTRIN 282.
FT REPEAT 7372 7396 SPECTRIN 283.
FT REPEAT 7396 7420 SPECTRIN 284.
FT REPEAT 7420 7444 SPECTRIN 285.
FT REPEAT 7444 7468 SPECTRIN 286.
FT REPEAT 7468 7492 SPECTRIN 287.
FT REPEAT 7492 7516 SPECTRIN 288.
FT REPEAT 7516 7540 SPECTRIN 289.
FT REPEAT 7540 7564 SPECTRIN 290.
FT REPEAT 7564 7588 SPECTRIN 291.
FT REPEAT 7588 7612 SPECTRIN 292.
FT REPEAT 7612 7636 SPECTRIN 293.
FT REPEAT 7636 7660 SPECTRIN 294.
FT REPEAT 7660 7684 SPECTRIN 295.
FT REPEAT 7684 7708 SPECTRIN 296.
FT REPEAT 7708 7732 SPECTRIN 297.
FT REPEAT 7732 7756 SPECTRIN 298.
FT REPEAT 7756 7780 SPECTRIN 299.
FT REPEAT 7780 7804 SPECTRIN 300.
FT REPEAT 7804 7828 SPECTRIN 301.
FT REPEAT 7828 7852 SPECTRIN 302.
FT REPEAT 7852 7876 SPECTRIN 303.
FT REPEAT 7876 7900 SPECTRIN 304.
FT REPEAT 7900 7924 SPECTRIN 305.
FT REPEAT 7924 7948 SPECTRIN 306.
FT REPEAT 7948 7972 SPECTRIN 307.
FT REPEAT 7972 7996 SPECTRIN 308.
FT REPEAT 7996 8020 SPECTRIN 309.
FT REPEAT 8020 8044 SPECTRIN 310.
FT REPEAT 8044 8068 SPECTRIN 311.
FT REPEAT 8068 8092 SPECTRIN 312.
FT REPEAT 8092 8116 SPECTRIN 313.
FT REPEAT 8116 8140 SPECTRIN 314.
FT REPEAT 8140 8164 SPECTRIN 315.
FT REPEAT 8164 8188 SPECTRIN 316.
FT REPEAT 8188 8212 SPECTRIN 317.
FT REPEAT 8212 8236 SPECTRIN 318.
FT REPEAT 8236 8260 SPECTRIN 319.
FT REPEAT 8260 8284 SPECTRIN 320.
FT REPEAT 8284 8308 SPECTRIN 321.
FT REPEAT 8308 8332 SPECTRIN 322.
FT REPEAT 8332 8356 SPECTRIN 323.
FT REPEAT 8356 8380 SPECTRIN 324.
FT REPEAT 8380 8404 SPECTRIN 325.
FT REPEAT 8404 8428 SPECTRIN 326.
FT REPEAT 8428 8452 SPECTRIN 327.
FT REPEAT 8452 8476 SPECTRIN 328.
FT REPEAT 8476 8500 SPECTRIN 329.
FT REPEAT 8500 8524 SPECTRIN 330.
FT REPEAT 8524 8548 SPECTRIN 331.
FT REPEAT 8548 8572 SPECTRIN 332.
FT REPEAT 8572 8596 SPECTRIN 333.
FT REPEAT 8596 8620 SPECTRIN 334.
FT REPEAT 8620 8644 SPECTRIN 335.
FT REPEAT 8644 8668 SPECTRIN 336.
FT REPEAT 8668 8692 SPECTRIN 337.
FT REPEAT 8692 8716 SPECTRIN 338.
FT REPEAT 8716 8740 SPECTRIN 339.
FT REPEAT 8740 8764 SPECTRIN 340.
FT REPEAT 8764 8788 SPECTRIN 341.
FT REPEAT 8788 8812 SPECTRIN 342.
FT REPEAT 8812 8836 SPECTRIN 343.
FT REPEAT 8836 8860 SPECTRIN 344.
FT REPEAT 8860 8884 SPECTRIN 345.
FT REPEAT 8884 8908 SPECTRIN 346.
FT REPEAT 8908 8932 SPECTRIN 347.
FT REPEAT 8932 8956 SPECTRIN 348.
FT REPEAT 8956 8980 SPECTRIN 349.
FT REPEAT 8980 9004 SPECTRIN 350.
FT REPEAT 9004 9028 SPECTRIN 351.
FT REPEAT 9028 9052 SPECTRIN 352.
FT REPEAT 9052 9076 SPECTRIN 353.
FT REPEAT 9076 9100 SPECTRIN 354.
FT REPEAT 9100 9124 SPECTRIN 355.
FT REPEAT 9124 9148 SPECTRIN 356.
FT REPEAT 9148 9172 SPECTRIN 357.
FT REPEAT 9172 9196 SPECTRIN 358.
FT REPEAT 9196 9220 SPECTRIN 359.
FT REPEAT 9220 9244 SPECTRIN 360.
FT REPEAT 9244 9268 SPECTRIN 361.
FT REPEAT 9268 9292 SPECTRIN 362.
FT REPEAT 9292 9316 SPECTRIN 363.
FT REPEAT 9316 9340 SPECTRIN 364.
FT REPEAT 9340 9364 SPECTRIN 365.
FT REPEAT 9364 9388 SPECTRIN 366.
FT REPEAT 9388 9412 SPECTRIN 367.
FT REPEAT 9412 9436 SPECTRIN 368.
FT REPEAT 9436 9460 SPECTRIN 369.
FT REPEAT 9460 9484 SPECTRIN 370.
FT REPEAT 9484 9508 SPECTRIN 371.
FT REPEAT 9508 9532 SPECTRIN 372.
FT REPEAT 9532 9556 SPECTRIN 373.
FT REPEAT 9556 9580 SPECTRIN 374.
FT REPEAT 9580 9604 SPECTRIN 375.
FT REPEAT 9604 9628 SPECTRIN 376.
FT REPEAT 9628 9652 SPECTRIN 377.
FT REPEAT 9652 9676 SPECTRIN 378.
FT REPEAT 9676 9700 SPECTRIN 379.
FT REPEAT 9700 9724 SPECTRIN 380.
FT REPEAT 9724 9748 SPECTRIN 381.
FT REPEAT 9748 9772 SPECTRIN 382.
FT REPEAT 9772 9796 SPECTRIN 383.
FT REPEAT 9796 9820 SPECTRIN 384.
FT REPEAT 9820 9844 SPECTRIN 385.
FT REPEAT 9844 9868 SPECTRIN 386.
FT REPEAT 9868 9892 SPECTRIN 387.
FT REPEAT 9892 9916 SPECTRIN 388.
FT REPEAT 9916 9940 SPECTRIN 389.
FT REPEAT 9940 9964 SPECTRIN 390.
FT REPEAT 9964 9988 SPECTRIN 391.
FT REPEAT 9988 10012 SPECTRIN 392.
FT REPEAT 10012 10036 SPECTRIN 393.
FT REPEAT 10036 10060 SPECTRIN 394.
FT REPEAT 10060 10084 SPECTRIN 395.
FT REPEAT 10084 10108 SPECTRIN 396.
FT REPEAT 10108 10132 SPECTRIN 397.
FT REPEAT 10132 10156 SPECTRIN 398.
FT REPEAT 10156 10180 SPECTRIN 399.
FT REPEAT 10180 10204 SPECTRIN 400.
FT REPEAT 10204 10228 SPECTRIN 401.
FT REPEAT 10228 10252 SPECTRIN 402.
FT REPEAT 10252 10276 SPECTRIN 403.
FT REPEAT 10276 10300 SPECTRIN 404.
FT REPEAT 10300 10324 SPECTRIN 405.
FT REPEAT 10324 10348 SPECTRIN 406.
FT REPEAT 10348 10372 SPECTRIN 407.
FT REPEAT 10372 10396 SPECTRIN 408.
FT REPEAT 10396 10420 SPECTRIN 409.
FT REPEAT 10420 10444 SPECTRIN 410.
FT REPEAT 10444 10468 SPECTRIN 411.
FT REPEAT 10468 10492 SPECTRIN 412.
FT REPEAT 10492 10516 SPECTRIN 413.
FT REPEAT 10516 10540 SPECTRIN 414.
FT REPEAT 10540 10564 SPECTRIN 415.
FT REPEAT 10564 10588 SPECTRIN 416.
FT REPEAT 10588 10612 SPECTRIN 417.
FT REPEAT 10612 10636 SPECTRIN 418.
FT REPEAT 10636 10660 SPECTRIN 419.
FT REPEAT 10660 10684 SPECTRIN 420.
FT REPEAT 10684 10708 SPECTRIN 421.
FT REPEAT 10708 10732 SPECTRIN 422.
FT REPEAT 10732 10756 SPECTRIN 423.
FT REPEAT 10756 10780 SPECTRIN 424.
FT REPEAT 10780 10804 SPECTRIN 425.
FT REPEAT 10804 10828 SPECTRIN 426.
FT REPEAT 10828 10852 SPECTRIN 427.
FT REPEAT 10852 10876 SPECTRIN 428.
FT REPEAT 10876 10900 SPECTRIN 429.
FT REPEAT 10900 10924 SPECTRIN 430.
FT REPEAT 10924 10948 SPECTRIN 431.
FT REPEAT 10948 10972 SPECTRIN 432.
FT REPEAT 10972 10996 SPECTRIN 433.
FT REPEAT 10996 11020 SPECTRIN 434.
FT REPEAT 11020 11044 SPECTRIN 435.
FT REPEAT 11044 11068 SPECTRIN 436.
FT REPEAT 11068 11092 SPECTRIN 437.
FT REPEAT 11092 11116 SPECTRIN 438.
FT REPEAT 11116 11140 SPECTRIN 439.
FT REPEAT 11140 11164 SPECTRIN 440.
FT REPEAT 11164 11188 SPECTRIN 441.
FT REPEAT 11188 11212 SPECTRIN 442.
FT REPEAT 11212 11236 SPECTRIN 443.
FT REPEAT 11236 11260 SPECTRIN 444.
FT REPEAT 11260 11284 SPECTRIN 445.
FT REPEAT 11284 11308 SPECTRIN 446.
FT REPEAT 11308 11332 SPECTRIN 447.
FT REPEAT 11332 11356 SPECTRIN 448.
FT REPEAT 11356 11380 SPECTRIN 449.
FT REPEAT 11380 11404 SPECTRIN 450.
FT REPEAT 11404 11428 SPECTRIN 451.
FT REPEAT 11428 11452 SPECTRIN 452.
FT REPEAT 11452 11476 SPECTRIN 453.
FT REPEAT 11476 11500 SPECTRIN 454.
FT REPEAT 11500 11524 SPECTRIN 455.
FT REPEAT 11524 11548 SPECTRIN 456.
FT REPEAT 11548 11572 SPECTRIN 457.
FT REPEAT 11572 11596 SPECTRIN 458.
FT REPEAT 11596 11620 SPECTRIN 459.
FT REPEAT 11620 11644 SPECTRIN 460.
FT REPEAT 11644 11668 SPECTRIN 461.
FT REPEAT 11668 11692 SPECTRIN 462.
FT REPEAT 11692 11716 SPECTRIN 463.
FT REPEAT 11716 11740 SPECTRIN 464.
FT REPEAT 11740 11764 SPECTRIN 465.
FT REPEAT 11764 11788 SPECTRIN 466.
FT REPEAT 11788 11812 SPECTRIN 467.
FT REPEAT 11812 11836 SPECTRIN 468.
FT REPEAT 11836 11860 SPECTRIN 469.
FT REPEAT 11860 11884 SPECTRIN 470.
FT REPEAT 11884 11908 SPECTRIN 471.
FT REPEAT 11908 11932 SPECTRIN 472.
FT REPEAT 11932 11956 SPECTRIN 473.
FT REPEAT 11956 11980 SPECTRIN 474.
FT REPEAT 11980 12004 SPECTRIN 475.
FT REPEAT 12004 12028 SPECTRIN 476.
FT REPEAT 12028 12052 SPECTRIN 477.
FT REPEAT 12052 12076 SPECTRIN 478.
FT REPEAT 12076 12100 SPECTRIN 479.
FT REPEAT 12100 12124 SPECTRIN 480.
FT REPEAT 12124 12148 SPECTRIN 481.
FT REPEAT 12148 12172 SPECTRIN 482.
FT REPEAT 12172 12196 SPECTRIN 483.
FT REPEAT 12196 12220 SPECTRIN 484.
FT REPEAT 12220 12244 SPECTRIN 485.
FT REPEAT 12244 12268 SPECTRIN 486.
FT REPEAT 12268 12292 SPECTRIN 487.
FT REPEAT 12292 12316 SPECTRIN 488.
FT REPEAT 12316 12340 SPECTRIN 489.
FT REPEAT 12340 12364 SPECTRIN 490.
FT REPEAT 12364 12388 SPECTRIN 491.
FT REPEAT 12388 12412 SPECTRIN 492.
FT REPEAT 12412 12436 SPECTRIN 493.
FT REPEAT 12436 12460 SPECTRIN 494.
FT REPEAT 12460 12484 SPECTRIN 495.
FT REPEAT 12484 12508 SPECTRIN 496.
FT REPEAT 12508 12532 SPECTRIN 497.
FT REPEAT 12532 12556 SPECTRIN 498.
FT REPEAT 12556 12580 SPECTRIN 499.
FT REPEAT 12580 12604 SPECTRIN 500.
FT REPEAT 12604 12628 SPECTRIN 501.
FT REPEAT 12628 12652 SPECTRIN 502.
FT REPEAT 12652 12676 SPECTRIN 503.
FT REPEAT 12676 12700 SPECTRIN 504.
FT REPEAT 12700 12724 SPECTRIN 505.
FT REPEAT 12724 12748 SPECTRIN
```


Db 1355 EVAENENDTNWTEHDWIAISL 1377

RESULT 3

E#9631

galactosamine-containing minor teichoic acid biosynthesis ggaB - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999

C:Accession: E9631

E:Kunst, P.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertezzo, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. E. Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallenbach, J.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetellelled, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, A.;Authors: Schleif, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, kenchu, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. A.;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A. A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis. A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: E9631

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-900 <KUN>

A:Cross-references: GB:Z99122; GB:AL009126; NID:g2636029; PID:CAB15585.1; PID:el184474;

A:Experimental source: strain 168

C:Genetics:

A:Gene: ggaB

```
Query Match      30.2%; Score 58; DB 2; Length 900;
Best Local Similarity 33.3%; Pred. No. 14;
Matches 11; Conservative 8; Mismatches 14; Indels 0; Gaps 0;
```

RESULT 4
T28409
ORF MSV248 probable inhibitor of apoptosis protein (IAP) - Melanoplus sanguinipes entomoc
C:Species: Melanoplus sanguinipes entomopoxvirus
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C:Accession: T28409
R:Afonso, C.J.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.
J. Virol. 73, 533-552, 1999
A:Title: The genome of Melanoplus sanguinipes entomopoxvirus.
A:Reference number: Z20484; MUID:99102612; PMID:9847359
A:Accession: T28409
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-150 <AFO>
A:Cross-references: EMBL:AF063866; NID:54049647; PIDN:AAC97724.1; PID:G4049764
C:Genetics:
A:Note: MSV248

```

Query Match      29.2%; Score 56; DB 2; Length 150;
Best Local Similarity 41.7%; Pred. No. 3.8;
Matches 10; Conservative 1; Mismatches 7; Indels 6; Gaps 1;

QY      7 KLINWAYQQVQQNKKEDAWIEHDVW 30
      ||| ||| ||| ||| ||| |||
Db      59 KIKNWLXY-----NDPWFIEHSKW 76

```

RESULT 5
S49113

hypothetical protein 2 - Microcystis aeruginosa
C;Species: Microcystis aeruginosa
C;Date: 01-Feb-1995 #sequence_revision 12-May-1995 #text_change 23-Mar-2001
C;Accession: S49113

R;Juerchott, K.; Boerner, T.
submitted to the EMBL Data Library, November 1993
A;Description: Sequence of the cyanobacterial plasmid pMA1 from Microcystis aeruginosa
A;Reference number: S49112
A;Accession: S49113
A>Status: preliminary
A:Molecule type: DNA
A;Residues: 1-502 <JUE>
A;Cross-references: EMBL:Z28337; NID:G509352; PIDN:CAA82191.1; PID:G509354
C;Superfamily: Microcystis aeruginosa hypothetical protein 2

Query Match 28.9%; Score 55.5; DB 2; Length 502;
Best Local Similarity 25.0%; Pred. No. 16;
Matches 12; Conservative 7; Mismatches 14; Indels 15; Gaps 1;

Qy 2 RSYRYKLINWA-----YOOVQNKKEDAWIEHDVWRMEI 34
:||||| : :||| :
Db 392 QAYREKLLGNWVRYLDPQLPEEFKALQAENKWNNDSWEDEL 439
:||||| : :||| :

RESULT 6
C69546
hypothetical protein AF2371 - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C;Accession: C69546
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glocke, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: C69546
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A;Residues: 1-152 <KLE>
A;Cross-references: GB:AE001112; GB:AE000782; NID:G2689435; PIDN:AAB91296.1; PID:G26507

Query Match 27.6%; Score 53; DB 2; Length 152;
Best Local Similarity 42.9%; Pred. No. 9.4;
Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 5 RYKLLNWAYOQQVNKKEDAMI 25
||||| : :||| :
Db 21 RYKLLSWLVDRYERNRGAYI 41
||||| : :||| :

RESULT 7
E69966
hypothetical protein yqkC - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C;Accession: E69966
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berthe
C.; Bron, S.; Bouilliet, S.; Bruschii, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
teich, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsleht, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinio
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel
Rieger, M.; Rivolta, C.; Rocha, Z.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Ser
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Togonoi, A.; Tosato, V.; Uchiya
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: E69966
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-79 <KUN>
A;Cross-references: GB:Z99116; GB:AL009126; NID:G2634723; PIDN:CAB14297.1; PID:e1185634;
A;Experimental source: strain 168
C;Genetics:
A;Gene: yqkC

Query Match 27.1%; Score 52; DB 2; Length 79;
Best Local Similarity 28.4%; Pred. No. 6.4;
Matches 9; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

Qy 2 RSYRYKLLNWAYQQVQQNKEDAWIEHDVWRME 33
Db 23 QNHPTLLHWSIGAESIKKDVWLLQDWTPE 54
::: |||::: :|||:|

RESULT 8
A41101
phorbol ester-binding protein unc-13 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change 23-Feb-1997
C;Accession: A41101
R;Maruyama, I.N.; Brenner, S.
Proc. Natl. Acad. Sci. U.S.A. 88, 5729-5733, 1991
A;Title: A phorbol ester/diacylglycerol-binding protein encoded by the unc-13 gene of *Caenorhabditis elegans*
A;Reference number: A41101; MUID:91288538; PMID:2062851
A;Accession: A41101
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1734 <WAR>
A;Cross-references: GB:M62830
C;Superfamily: protein kinase C zinc-binding repeat homology
C;Keywords: phosphoprotein
F;615-664/Domain: protein kinase C zinc-binding repeat homology <KZ2>

Query Match 27.1%; Score 52; DB 2; Length 1734;
Best Local Similarity 40.0%; Pred. No. 1.7e+02;
Matches 8; Conservative 5; Mismatches 5; Indels 2; Gaps 1;

Qy 10 NWAYQQVQQ--NKKEDAWIEH 27
Db 430 NWRVDSIQEDNEKDNWKQH 449
|||::: :|||:|

RESULT 9
T19295
hypothetical protein ZK524.2a - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Sep-2000
C;Accession: T19295; T19331; T27889
R;Gardner, A.
submitted to the EMBL Data Library, September 1996
A;Reference number: Z19103
A;Accession: T19295
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-1813 <WIL>
A;Cross-references: EMBL:Z79694; PIDN:CAB01966.1; GSPDB:GN00019; CESP:ZK524.2a
A;Experimental source: clone C15A11
R;Lloyd, C.
submitted to the EMBL Data Library, March 1997
A;Reference number: Z19198
A;Accession: T19931
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-1813 <W12>
A;Cross-references: EMBL:Z92779; PIDN:CAB07173.1; GSPDB:GN00019; CESP:ZK524.2a
A;Experimental source: clone C44E1
R;Gardner, A.

submitted to the EMBL Data Library, June 1996
A;Reference number: Z20435
A;Accession: T27889
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-1813 <W13>
A;Cross-references: EMBL:Z73912; PIDN:CAA98147.1; GSPDB:GN00019; CESP:ZK524.2a
A;Experimental source: clone ZK524
C;Genetics:
A;Gene: CESP:ZK524.2a
A;Map position: 1
A;Introns: 25/1; 35/1; 68/2; 105/2; 149/1; 170/3; 186/1; 235/3; 412/2; 434/2; 500/2; 540/1; 1781/3
C;Superfamily: protein kinase C zinc-binding repeat homology
F;693-742/Domain: protein kinase C zinc-binding repeat homology <KZN>

Query Match 27.1%; Score 52; DB 2; Length 1813;
Best Local Similarity 40.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 5; Mismatches 5; Indels 2; Gaps 1;

Qy 10 NWAYQQVQQ--NKKEDAWIEH 27
Db 511 NWRVDSIQEDNEKDNWKQH 530
|||::: :|||:|

RESULT 10
G64606
hydantoin utilization protein A - *Helicobacter pylori* (strain 26695)
C;Species: *Helicobacter pylori*
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 16-Feb-2001
C;Accession: G64606
R;Tomb, J.P.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenna, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Nature 388, 539-547, 1997
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A;Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: G64606
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-713 <TOM>
A;Cross-references: GB:AE000582; GB:AE000511; NID:G2313812; PIDN:AAD07746.1; PID:G23138;
C;Superfamily: Pseudomonas D-amino acid hydantoin hydrolase (ATP-hydrolyzing) hyd

Query Match 26.8%; Score 51.5; DB 2; Length 713;
Best Local Similarity 30.3%; Pred. No. 77;
Matches 10; Conservative 5; Mismatches 3; Indels 15; Gaps 1;

Qy 1 RRSYRYKLLNWAYQQVQQNKEDAWIEHDVWRME 33
Db 653 RKFYRHK-----KWVDADVMQME 670
|||::: :|||:|

RESULT 11
AB2102
hypothetical protein asr2369 [imported] - *Nostoc* sp. (strain PCC 7120)
C;Species: *Nostoc* sp. PCC 7120
A;Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AB2102
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, M.; DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena* sp. strain PCC 7120
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AB2102
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-90 <KUR>
A;Cross-references: GB:BA000019; PIDN:BA074068.1; PID:g17131461; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:10:33 ; Search time 60.093 Seconds
(without alignments)
176.480 Million cell updates/sec

Title: US-10-010-667A-2_COPY_185_218

Perfect score: 192

Sequence: 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1279676 seqs, 311918243 residues

Total number of hits satisfying chosen parameters: 1279676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

```
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	192	100.0	34	US-10-011-095-20
2	192	100.0	34	Sequence 20, Appl
3	192	100.0	34	Sequence 20, Appl
4	192	100.0	34	Sequence 38, Appl
5	192	100.0	267	US-09-747-835A-50
6	192	100.0	267	US-09-747-835A-51
7	192	100.0	267	US-10-312-312-50
8	192	100.0	267	US-10-312-312-51
9	192	100.0	339	US-09-758-143-879
10	192	100.0	339	US-09-780-669-879
11	192	100.0	339	US-09-822-827-879
12	192	100.0	339	US-09-802-520-11
13	192	100.0	339	US-09-895-793-879
14	192	100.0	339	US-09-895-814-879
15	192	100.0	339	US-10-205-267-13
			339	US-10-408-009-2

16	192	100.0	339	13	US-10-012-896-879	Sequence 879, App
17	192	100.0	339	14	US-10-011-095-2	Sequence 2, Appli
18	192	100.0	339	14	US-10-010-667A-2	Sequence 2, Appli
19	192	100.0	339	14	US-10-205-823-397	Sequence 397, App
20	192	100.0	339	14	US-10-144-678A-879	Sequence 879, App
21	192	100.0	339	14	US-10-294-025-879	Sequence 879, App
22	192	100.0	339	15	US-10-239-607-37	Sequence 37, Appl
23	192	100.0	339	15	US-10-295-027-714	Sequence 714, App
24	192	100.0	339	15	US-10-295-027-1347	Sequence 1347, Ap
25	192	100.0	368	12	US-10-425-114-72779	Sequence 72779, A
26	192	100.0	375	14	US-10-165-044-2	Sequence 2, Appli
27	110	57.3	443	12	US-10-455-822-93	Sequence 93, Appl
28	107	55.7	141	9	US-09-963-896-1	Sequence 1, Appli
29	107	55.7	173	14	US-10-011-095-8	Sequence 8, Appli
30	107	55.7	173	14	US-10-010-667A-8	Sequence 8, Appli
31	107	55.7	237	9	US-09-747-835A-15	Sequence 15, Appl
32	107	55.7	237	12	US-10-312-312-15	Sequence 15, Appl
33	107	55.7	419	12	US-10-455-822-11	Sequence 11, Appl
34	107	55.7	419	12	US-10-455-822-80	Sequence 80, Appl
35	107	55.7	419	12	US-10-455-822-172	Sequence 172, App
36	107	55.7	419	12	US-10-455-822-174	Sequence 174, App
37	107	55.7	419	12	US-10-455-822-195	Sequence 195, App
38	107	55.7	419	15	US-10-239-607-32	Sequence 32, Appl
39	107	55.7	444	12	US-10-455-822-88	Sequence 88, Appl
40	107	55.7	444	12	US-10-455-822-89	Sequence 89, Appl
41	107	55.7	444	12	US-10-455-822-90	Sequence 90, Appl
42	107	55.7	444	12	US-10-455-822-91	Sequence 91, Appl
43	107	55.7	444	12	US-10-455-822-92	Sequence 92, Appl
44	107	55.7	454	10	US-09-888-257A-10	Sequence 10, Appl
45	107	55.7	454	12	US-10-455-822-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1

```
US-10-011-095-20
; Sequence 20, Application US/10011095
; Publication No. US20030045682A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Hubert, Rene S.
; APPLICANT: Leong, Kahan
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Safran, Douglas C.
; APPLICANT: Mitchell, Steve Chappell
; TITLE OF INVENTION: ANTIBODIES IMMUNOSPECIFIC FOR STEAP1 (AS AMENDED)
; FILE REFERENCE: 511582001610
; CURRENT APPLICATION NUMBER: US/10/011.095
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: STEAP-1 peptide
US-10-011-095-20
```

Query Match 100.0%; Score 192; DB 14; Length 34;

Best Local Similarity 100.0%; Pred. No. 1.8e-18;

Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34

Db 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34

```
RESULT 2
US-10-010-667A-20
; Sequence 20, Application US/10010667A
; Publication No. US20030055217A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Hubert, Rene S.
; APPLICANT: Leong, Kahan
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Saffran, Douglas C.
; APPLICANT: Mitchell, Steve Chappell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 511582001601
; CURRENT APPLICATION NUMBER: US/10/010,667A
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: STEAP-1 peptide
US-10-010-667A-20

Query Match          100.0%; Score 192; DB 14; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.8e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSYRYKLLNWAYQQVQQNKEDAWIEHDVWRMEI 34
Db 1 RRSYRYKLLNWAYQQVQQNKEDAWIEHDVWRMEI 34

RESULT 3
US-10-165-044-38
; Sequence 38, Application US/10165044
; Publication No. US20030149531A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas Saffran
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Steven Chappell Mitchell
; APPLICANT: Mary Faris
; APPLICANT: Ava Jakobovits
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 51158-20016.02
; CURRENT APPLICATION NUMBER: US/10/165,044
; CURRENT FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: US 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: US 60/091,183
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 09/455,486
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: WO 99/62941
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/33040
; PRIOR FILING DATE: 2000-12-06
```

```
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-165-044-38

Query Match          100.0%; Score 192; DB 14; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.8e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSYRYKLLNWAYQQVQQNKEDAWIEHDVWRMEI 34
Db 1 RRSYRYKLLNWAYQQVQQNKEDAWIEHDVWRMEI 34

RESULT 4
US-09-747-835A-50
; Sequence 50, Application US/09747835A
; Patent No. US20020146692A1
; GENERAL INFORMATION:
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Drmauac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
; TITLE OF INVENTION: LIKE) POLYPEPTIDES AND POLYNUCLEOTIDES
; FILE REFERENCE: HYS-37CIP
; CURRENT APPLICATION NUMBER: US/09/747,835A
; CURRENT FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US 09/729,739
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US 09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-747-835A-50

Query Match          100.0%; Score 192; DB 9; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.6e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSYRYKLLNWAYQQVQQNKEDAWIEHDVWRMEI 34
Db 119 RRSYRYKLLNWAYQQVQQNKEDAWIEHDVWRMEI 152

RESULT 5
US-09-747-835A-51
; Sequence 51, Application US/09747835A
; Patent No. US20020146692A1
; GENERAL INFORMATION:
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
```

```
; APPLICANT: Zhou, Ping
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
; TITLE OF INVENTION: (LIKE) POLYPEPTIDES AND POLYNUCLEOTIDES
; FILE REFERENCE: HYS-37CIP
; CURRENT APPLICATION NUMBER: US/09/747,835A
; CURRENT FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US 09/729,739
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US 09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 51
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-09-747-835A-51
```

```
Query Match 100.0%; Score 192; DB 9; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.6e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 34
Db 119 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 152
```

RESULT 6

```
US-10-312-312-50
; Sequence 50, Application US/10312312
; Publication No. US20040068097A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
; TITLE OF INVENTION: (LIKE) POLYPEPTIDES AND POLYNUCLEOTIDES
; FILE REFERENCE: 21272-015-061/HYS-37CIP
; CURRENT APPLICATION NUMBER: US/10/312,312
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 09/729,739
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US 09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 63
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-10-312-312-50
Query Match 100.0%; Score 192; DB 12; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.6e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 34
Db 119 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 152
```

RESULT 7

```
US-10-312-312-51
; Sequence 51, Application US/10312312
; Publication No. US20040068097A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
; TITLE OF INVENTION: (LIKE) POLYPEPTIDES AND POLYNUCLEOTIDES
; FILE REFERENCE: 21272-015-061/HYS-37CIP
; CURRENT APPLICATION NUMBER: US/10/312,312
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 09/729,739
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US 09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 51
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-10-312-312-51
```

```
Query Match 100.0%; Score 192; DB 12; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.6e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 34
Db 119 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 152
```

RESULT 8

```
US-09-759-143-879
; Sequence 879, Application US/09759143
; Patent No. US2002002248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
```

APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqui
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C23
CURRENT APPLICATION NUMBER: US/09/759,143
CURRENT FILING DATE: 2001-01-12
NUMBER OF SEQ ID NOS: 934
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 879
LENGTH: 339
TYPE: PRT
ORGANISM: Homo sapiens
US-09-759-143-879

Query Match 100.0%; Score 192; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34
|||||
DB 185 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 218

RESULT 9

US-09-780-669-879
Sequence 879, Application US/09780669
Patent No. US20020051977A1
GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqui
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C24
CURRENT APPLICATION NUMBER: US/09/780,669
CURRENT FILING DATE: 2001-02-09
NUMBER OF SEQ ID NOS: 943
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 879
LENGTH: 339
TYPE: PRT
ORGANISM: Homo sapiens
US-09-780-669-879

Query Match 100.0%; Score 192; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34
|||||
DB 185 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 218

RESULT 12

US-09-895-793-879

Query Match 100.0%; Score 192; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34
|||||
DB 185 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 218

RESULT 10

US-09-822-827-879
Sequence 879, Application US/09822827
Patent No. US20020081680A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.534C1
CURRENT APPLICATION NUMBER: US/09/822,827
CURRENT FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 982
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 879
LENGTH: 339
TYPE: PRT
ORGANISM: Homo sapiens
US-09-822-827-879

Query Match 100.0%; Score 192; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34
|||||
DB 185 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 218

RESULT 11

US-09-802-520-11
Sequence 11, Application US/09802520
Publication No. US20020187472A1
GENERAL INFORMATION:

APPLICANT: Lal, Preeti
APPLICANT: Paris, Mary
APPLICANT: Chen, Huei-Mei
APPLICANT: Ison, Craig H.
TITLE OF INVENTION: STEAP-RELATED PROTEIN
FILE REFERENCE: PC-0037 US
CURRENT APPLICATION NUMBER: US/09/802,520
CURRENT FILING DATE: 2001-03-09
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PERL Program
SEQ ID NO 11
LENGTH: 339
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20020187472A1 g6572948
US-09-802-520-11

Query Match 100.0%; Score 192; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34
|||||
DB 185 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 218

; Sequence 879, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriack
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121-534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-793-879

Query Match 100.0%; Score 192; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 34
|||||
Db 185 RRSRYRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 218

RESULT 13
US-09-895-814-879
; Sequence 879, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriack
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121-427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-814-879

Query Match 100.0%; Score 192; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 34
|||||
Db 185 RRSRYRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 218

RESULT 14
US-10-205-267-13
; Sequence 13, Application US/10205267
; Publication No. US20030064397A1
; GENERAL INFORMATION:
; APPLICANT: Spaccake, Kimberly M.
; APPLICANT: Rickert Paula K.
; APPLICANT: Lal, Preeti G.
; APPLICANT: Ison, Craig H.
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN PROSTATE AND LI
; TITLE OF INVENTION: TUMORS
; FILE REFERENCE: PV-0008 CIP
; CURRENT APPLICATION NUMBER: US/10/205,267
; CURRENT FILING DATE: 2002-07-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 339
; TYPE: PRT
; ORGANISM:
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Genbank ID No: 96572948
US-10-205-267-13

Query Match 100.0%; Score 192; DB 12; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 34
|||||
Db 185 RRSRYRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 218

RESULT 15
US-10-408-009-2
; Sequence 2, Application US/10408009
; Publication No. US20040072196A1
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Stephen C. Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 511582001603
; CURRENT APPLICATION NUMBER: US/10/408,009
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 09/455,486
; PRIOR FILING DATE: 1999-12-06

```
; PRIOR APPLICATION NUMBER: 09/323,873
; PRIOR FILING DATE: 1993-06-01
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-009-2

Query Match      100.0%; Score 192; DB 12; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RRSYRYKLLNWAYQQVQQNKEDAWIEHDVWRMEI 34
      |||||||||||||||||||||||||||||||||||
Db      185 RRSYRYKLLNWAYQQVQQNKEDAWIEHDVWRMEI 218

Search completed: July 12, 2004, 14:27:49
Job time : 60.093 secs
```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:04:08 ; Search time 20.9535 Seconds
(without alignments)
83.770 Million cell updates/sec

Title: US-10-010-667A-2_COPY_185_218
Perfect score: 192
Sequence: 1 RRSRYKLLNWAYQQVQNKEDAWIEHDVVRMEI 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTCUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	192	100.0	34	4	US-09-323-873A-20
2	192	100.0	339	4	US-09-323-873A-2
3	192	100.0	339	4	US-09-685-166A-879
4	107	55.7	141	3	US-09-083-521-1
5	107	55.7	173	4	US-09-323-873A-8
6	85	44.3	15	4	US-09-323-873A-32
7	57	29.7	320	4	US-09-543-681A-5163
8	52.5	27.3	340	4	US-09-543-681A-7850
9	52	27.1	166	4	US-09-328-352-4565
10	51	26.6	789	4	US-09-252-991A-28767
11	50	26.0	370	4	US-09-149-476-436
12	49.5	25.8	940	4	US-09-198-452A-111
13	49	25.5	30	4	US-09-205-258-646
14	49	25.5	120	4	US-09-205-258-644
15	48.5	25.3	654	3	US-08-560-005-10
16	48.5	25.3	654	3	US-09-418-540-10
17	48.5	25.3	654	4	US-09-969-528-10
18	48	25.0	405	4	US-09-134-000C-5472
19	48	25.0	2037	4	US-09-543-681A-5538
20	47.5	24.7	178	4	US-09-134-001C-3412
21	47.5	24.7	4150	3	US-09-428-517-2
22	47	24.5	258	4	US-09-107-532A-6289
23	47	24.5	644	4	US-09-198-452A-63
24	47	24.5	1503	4	US-09-600-087-2
25	46.5	24.2	379	4	US-09-252-991A-26357
26	46.5	24.2	602	4	US-09-489-039A-12436
27	46.5	24.2	927	4	US-09-540-236-2607

28	46	24.0	111	4	US-09-107-532A-5449	Sequence 5449, Ap
29	46	24.0	236	3	US-09-121-973-4	Sequence 4, Appli
30	46	24.0	236	3	US-09-332-319-4	Sequence 4, Appli
31	46	24.0	236	4	US-09-239-867-2	Sequence 2, Appli
32	46	24.0	477	4	US-09-198-452A-19	Sequence 19, Appl
33	46	24.0	599	4	US-09-489-039A-13050	Sequence 13050, A
34	46	24.0	605	4	US-09-252-991A-26916	Sequence 26916, A
35	46	24.0	961	4	US-09-328-352-7358	Sequence 7358, Ap
36	45.5	23.7	441	4	US-09-107-532A-4173	Sequence 4173, Ap
37	45.5	23.7	605	4	US-09-252-991A-20980	Sequence 20980, A
38	45.5	23.7	854	4	US-09-206-551-17	Sequence 17, Appl
39	45	23.4	280	4	US-09-489-039A-9395	Sequence 9395, Ap
40	45	23.4	371	1	US-08-746-789A-2	Sequence 2, Appli
41	45	23.4	371	4	US-09-570-593-5	Sequence 5, Appli
42	45	23.4	516	4	US-09-489-039A-9477	Sequence 9477, Ap
43	45	23.4	526	4	US-09-198-452A-67	Sequence 67, Appl
44	45	23.4	631	4	US-09-107-532A-6640	Sequence 6640, Ap
45	45	23.4	837	4	US-09-489-039A-9385	Sequence 9385, Ap

ALIGNMENTS

RESULT 1
US-09-323-873A-20
; Sequence 20, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Kahan Leong
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 129.16USU2
; CURRENT APPLICATION NUMBER: US/09/323,873A
; CURRENT FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: STEAP-1 PEPTIDE
US-09-323-873A-20
Query Match 100.0%; Score 192; DB 4; Length 34;
Best Local Similarity 100.0%; Pred. No. 5.5e-20;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 RRSRYKLLNWAYQQVQNKEDAWIEHDVVRMEI 34
Db 1 RRSRYKLLNWAYQQVQNKEDAWIEHDVVRMEI 34
RESULT 2
US-09-323-873A-2
; Sequence 2, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Kahan Leong
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell

; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 129.16USU2
; CURRENT APPLICATION NUMBER: US/09/323,873A
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-323-873A-2

Query Match 100.0%; Score 192; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 7.2e-19;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQNKEDAWIEHDVWRMEI 34
(((|||||))))
Db 185 RRSRYKLLNWAYQQVQNKEDAWIEHDVWRMEI 218

RESULT 3

US-09-685-166A-879
; Sequence 879, Application US/09685166A
; Patent No. 6630305

; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-685-166A-879

Query Match 100.0%; Score 192; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 7.2e-19;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQNKEDAWIEHDVWRMEI 34
(((|||||))))
Db 185 RRSRYKLLNWAYQQVQNKEDAWIEHDVWRMEI 218

RESULT 4

US-09-083-521-1

; Sequence 1, Application US/09083521

; Patent No. 6048970
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: PROSTATE GROWTH-ASSOCIATED MEMBRANE PROTEINS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,521
; FILING DATE: Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0527 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 141 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSTUT10
; CLONE: 1691243
US-09-083-521-1

Query Match 55.7%; Score 107; DB 3; Length 141;
Best Local Similarity 58.8%; Pred. No. 1.5e-07;
Matches 20; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQNKEDAWIEHDVWRMEI 34
|||||
Db 13 RRSRYKLLNWAYQQVQHANIENSWNEEVWRMEI 46

RESULT 5

US-09-323-873A-8

; Sequence 8, Application US/09323873A
; Patent No. 6329503

; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar

; APPLICANT: Rene S. Hubert
; APPLICANT: Kahan Leong
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran

; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 129.16USU2
; CURRENT APPLICATION NUMBER: US/09/323,873A
; CURRENT FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8

; LENGTH: 173
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-323-873A-8

Query Match 55.7%; Score 107; DB 4; Length 173;
Best Local Similarity 58.8%; Pred. No. 1.9e-07;
Matches 20; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 RRSYRYKLNWAYQQVQKQKEDAWIEHDVVRMEI 34
Db 81 RRSERYFLNWAYQQVQKQKEDAWIEHDVVRMEI 114

RESULT 6

US-09-323-873A-32
; Sequence 32, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Kahan Leong
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 129.16USU2
; CURRENT APPLICATION NUMBER: US/09/323,873A
; PRIOR FILING DATE: 1999-06-01
; PRIOR FILING DATE: 1998-06-01
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo Sapiens

US-09-323-873A-32

Query Match 44.3%; Score 85; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 YQVQVQKQKEDAWIEH 27
Db 1 YQVQVQKQKEDAWIEH 15

RESULT 7

US-09-543-681A-5163
; Sequence 5163, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5163
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Proteus mirabilis

US-09-543-681A-5163

Query Match 29.7%; Score 57; DB 4; Length 320;
Best Local Similarity 38.7%; Pred. No. 3.2;

Matches 12; Conservative 5; Mismatches 8; Indels 6; Gaps 2;

Qy 6 YKLLN-WAYQQVQKQKED-----AWIEHDVW 30
Db 78 YAVMNEWAYNEYNQKDKDGLSAAWGVNDRW 108

RESULT 8

US-09-543-681A-7850
; Sequence 7850, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7850
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7850

Query Match 27.3%; Score 52.5; DB 4; Length 340;
Best Local Similarity 30.3%; Pred. No. 14;
Matches 10; Conservative 8; Mismatches 8; Indels 7; Gaps 1;

Qy 3 SYRY-----KLLNWAYQQVQKQKEDAWIEHD 28
Db 169 THYTAGTTDIALKRAYQQMKTLEEWLKR 201

RESULT 9

US-09-328-352-4565
; Sequence 4565, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4565
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4565

Query Match 27.1%; Score 52; DB 4; Length 166;
Best Local Similarity 27.6%; Pred. No. 7.5;
Matches 8; Conservative 7; Mismatches 10; Indels 4; Gaps 1;

Qy 4 YRYK-----LLNWAYQQVQKQKEDAWIEHD 28
Db 133 YEYNEKGVLLWTHHDPQNRHENGWLKHN 161

RESULT 10

US-09-252-991A-28767
; Sequence 28767, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A

```
/ CURRENT FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 28767
/ LENGTH: 789
/ TYPE: PRT
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28767

Query Match      26.6%; Score 51; DB 4; Length 789;
Best Local Similarity 52.9%; Pred. No. 59;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY      16 VQONKEDAWIEHDVWRM 32
Db      397 VORSKDDAVRHFWRLL 413

RESULT 11
US-09-149-476-436
/ Sequence 436, Application US/09149476
/ Patent No. 6420526
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: 186 Human Secreted proteins
/ CURRENT APPLICATION NUMBER: US/09/149,476
/ CURRENT FILING DATE: 1998-09-08
/ EARLIER APPLICATION NUMBER: PCT/US98/04493
/ EARLIER FILING DATE: 1998-03-06
/ EARLIER APPLICATION NUMBER: 60/040,162
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/040,333
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/038,621
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/040,626
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/040,334
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/040,336
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/040,163
/ EARLIER FILING DATE: 1997-03-07
/ EARLIER APPLICATION NUMBER: 60/047,600
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,615
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,597
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,502
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,633
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,583
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,617
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,618
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,503
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,592
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,581
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,584
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,500
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,587
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,492
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,598
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,613
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,582
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,596
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,612
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,632
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,601
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/043,580
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,568
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,314
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,569
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,311
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,671
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,674
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,669
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,312
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,313
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,672
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,315
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/048,974
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/056,886
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,877
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,889
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,893
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,630
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,878
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,662
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,872
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,882
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,637
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,903
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,888
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,879
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,880
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,894
```

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05

EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 26.0%; Score 50; DB 4; Length 370;
Best Local Similarity 40.7%; Pred. No. 35;
Matches 11; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 5 RYKLNWAYQQVQONKEDAWIEHDVWR 31
| : ||| : : : : : |||
Db 191 RQKLNDALELKEAFNEDFDVWR 217

RESULT 12

US-09-198-452A-111
Sequence 111, Application US/09198452A

Patent No. 6559294

GENERAL INFORMATION:

APPLICANT: Griffiths, R.

TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention

TITLE OF INVENTION: and treatment of infection

FILE REFERENCE: 9710-003-999

CURRENT APPLICATION NUMBER: US/09/198,452A

CURRENT FILING DATE: 1998-11-24

NUMBER OF SEQ ID NOS: 6849

SEQ ID NO 111

LENGTH: 940

TYPE: PRT

ORGANISM: Chlamydia pneumoniae

US-09-198-452A-111

Query Match 25.8%; Score 49.5; DB 4; Length 940;
Best Local Similarity 31.6%; Pred. No. 1.2e+02;
Matches 12; Conservative 8; Mismatches 9; Indels 9; Gaps 2;

QY 4 YR-YKLNW-----AYQVQONKEDAWIEHDVWR 32
||| : : : : : |||
Db 172 YRGYLVNWDPLQTLADDEVEYEXDGLWYIYR 209

RESULT 13

US-09-205-258-646

Sequence 646, Application US/09205258

Patent No. 6525174

GENERAL INFORMATION:

APPLICANT: Young et al.

TITLE OF INVENTION: 207 Human Secreted Proteins

FILE REFERENCE: P2007F1

CURRENT APPLICATION NUMBER: US/09/205,258

CURRENT FILING DATE: 1998-12-04

EARLIER APPLICATION NUMBER: PCT/US98/11422

EARLIER FILING DATE: 1998-06-04

EARLIER APPLICATION NUMBER: 60/048,885

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,375

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,881

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,880

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,896

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,020

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,876

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,895

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,884

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,894

EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,971
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,882
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SQ ID NO 646
LENGTH: 30
TYPE: PRT
ORGANISM: Homo sapiens
US-09-205-258-646

Query Match 25.58; Score 49; DB 4; Length 30;
Best Local Similarity 38.1%; Pred. No. 2.9;
Matches 8; Conservative 5; Mismatches 6; Indels 2; Gaps 1;

QY 9 INWAYQ--QVQNKEDAWIWH 27
: ||:: ||||: : |

Db 7 VFWAFESLQVQNRPERWASH 27
RESULT 14
US-09-205-258-644
; Sequence 644, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06


```
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 644
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (105)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (115)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;
US-09-205-258-644

Query Match          25.5%; Score 49; DB 4; Length 120;
Best Local Similarity 38.1%; Pred. No. 14;
Matches 8; Conservative 5; Mismatches 6; Indels 2; Gaps 1;

Qy 9 LNWAYQ--VQONKEDAWIEH 27
   :||:  |||:  |||:  |||:  |||:  |||:  |||:  |||:  |||:  |||:
Db 79 VRWAFESLQVPQNRPERWASH 99

RESULT 15
US-08-560-005-10
; Sequence 10, Application US/08560005
; Patent No. 6001354
; GENERAL INFORMATION:
; APPLICANT: Pot, David A.
; APPLICANT: Williams, Lewis T.
; APPLICANT: Jefferson, Anne Bennett
; APPLICANT: Majerus, Philip W.
; TITLE OF INVENTION: No. 6001354el Grb2 Associating Protein and Nucleic
; TITLE OF INVENTION: Acids Encoding Therefor
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,005
; FILING DATE:
; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-0624000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 654 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..654
; OTHER INFORMATION: /note= "consensus"
;
US-08-560-005-10

Query Match          25.3%; Score 48.5; DB 3; Length 654;
Best Local Similarity 25.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 12; Mismatches 9; Indels 3; Gaps 1;

Qy 3 SVRYKLLNWAYQQ---VQONKEDAWIEHDVWR 31
   :||:  |||:  |||:  |||:  |||:  |||:  |||:  |||:  |||:  |||:
Db 348 TYKRGDYAYKQKDTSEKTRVPAWCDRLMK 379

Search completed: July 12, 2004, 14:12:24
Job time : 21.9535 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 13:51:07 ; Search time 74.7209 Seconds

(without alignments)
128.567 Million cell updates/sec

Title: US-10-010-667A-2_COPY_185_218

Perfect score: 192

Sequence: 1 RRSYRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A Geneseq_29Jan04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	192	100.0	34	3	AAY58199 Human STR
2	192	100.0	34	4	AAE02787 Extracell
3	192	100.0	129	4	AAE02787 Extracell
4	192	100.0	129	4	AAE02787 Extracell
5	192	100.0	254	6	ABU98426 STEAP-1 v
6	192	100.0	254	6	ABU98426 STEAP-1 v
7	192	100.0	255	6	ABU98425 STEAP-1 v
8	192	100.0	255	6	ABU98428 STEAP-1 v
9	192	100.0	258	6	ABU98395 Novel hum
10	192	100.0	258	6	ABU98389 Novel hum
11	192	100.0	258	6	ABU98424 STEAP-1 v
12	192	100.0	258	6	ABU98392 Novel hum
13	192	100.0	258	6	ABU98387 Novel hum
14	192	100.0	258	6	ABU98393 Novel hum
15	192	100.0	258	6	ABU98390 Novel hum
16	192	100.0	258	6	ABU98391 Novel hum
17	192	100.0	258	6	ABU98384 Novel hum
18	192	100.0	258	6	ABU98431 STEAP-1 v
19	192	100.0	258	6	ABU98396 Novel hum
20	192	100.0	258	6	ABU98397 Novel hum
21	192	100.0	258	6	ABU98394 Novel hum
22	192	100.0	258	6	ABU98388 Novel hum
23	192	100.0	258	6	ABU98386 Novel hum
24	192	100.0	258	6	ABU98399 Novel hum
25	192	100.0	258	6	ABU98398 Novel hum

26	192	100.0	267	6	ABU60886 Human G p
27	192	100.0	267	6	ABU60887 Human G p
28	192	100.0	282	6	ABU98432 STEAP-1 v
29	192	100.0	282	6	ABU98427 STEAP-1 v
30	192	100.0	282	6	ABU98385 Novel hum
31	192	100.0	339	3	AAV58194 Human STR
32	192	100.0	339	4	AAU01282 P789P ami
33	192	100.0	339	4	AAU69927 Human pro
34	192	100.0	339	4	AAU78845 Human pro
35	192	100.0	339	4	AAU78845 Human pro
36	192	100.0	339	5	ABU95387 Human P78
37	192	100.0	339	5	ABU61813 Prostate
38	192	100.0	339	5	ABU98383 Novel hum
39	192	100.0	339	6	ABU98414 STEAP-1 v
40	192	100.0	339	6	ABU98430 STEAP-1 v
41	192	100.0	339	6	ABU98449 Prostate
42	192	100.0	339	7	ABU63313 Human six
43	192	100.0	339	7	ADB75573 Prostate
44	192	100.0	339	7	ADB14329 Human pro
45	192	100.0	375	4	AAE02780 Human six

ALIGNMENTS

RESULT 1

AAV58199
ID AAY58199 standard; peptide; 34 AA.
XX
AC AAY58199;
XX
DT 14-MAR-2000 (first entry)
XX
DE Human STRAP-1 peptide, corresponding to STRAP-1 extracellular region 2.
XX
KW Serpentine transmembrane antigen of the prostate; STRAP-1; prostate;
transmembrane domain; type IIa membrane protein; expression; cancer;
prostate cancer; bladder cancer; colon cancer; pancreatic cancer;
ovarian cancer; tumour antigen; immunisation; immune response; cellular;
humoral; anticancer vaccine; antibody; detection; diagnosis; prognosis;
monitoring; susceptibility; therapeutic inhibitor; drug targeting;
recombinant protein.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9962941-A2.
XX
PD 09-DEC-1999.
XX
PF 01-JUN-1999; 99WO-US012157.
XX
PR 01-JUN-1998; 98US-0087520P.
PR 30-JUN-1998; 98US-0091183P.
XX
PA (UROC-) UROGENESYS INC.
PA (AFAR/) AFAR D E.
PA (HUBE/) HUBERT R S.
PA (LEON/) LEONG K.
PA (RAIT/) RAITANO A B.
PA (SAFF/) SAFFRAN D C.
XX
PI Afar DE, Hubert RS, Leong K, Raitano AB, Safran DC;
XX
DR WPI; 2000-072832/06.
XX
PT Novel proteins useful as diagnostic markers and therapeutic targets,
particularly for prostatic cancer.
XX
PS Disclosure; Page 22; 83pp; English.
XX
CC Sequences AAY58198-Y58200 represent synthetic peptides that correspond to
the extracellular regions of STRAP-1 (serpentine transmembrane antigen of

the prostate, AAY58194). These peptides were used to raise monoclonal anti-STRAP-1 antibodies. STRAP-1 is the prototype member of the STRAP family of proteins (AAY58194-Y58197) which exhibit a high degree of structural conservation, but which show no significant structural homology to known human proteins. The STRAP-1 gene has been localised to chromosome 7p22. STRAP-1 is thought to be a type IIIa membrane protein and is expressed predominantly in prostate cells in normal human tissues. Structurally, STRAP-1 is a 339 amino acid protein characterised by six transmembrane domains and intracellular N- and C-termini, suggesting that it folds in a "serpentine" manner into three extracellular and two intracellular loops. STRAP-1 mRNA and protein expression is maintained at high levels and throughout all stages of prostate cancer. STRAP-1 mRNA and/or protein is also overexpressed in certain other cancers, including bladder, colon, pancreatic and ovarian cancer. The function of the STRAP proteins is not known. They may be ion channels (from the presence of six transmembrane domains, a feature which is shared by certain ion channels) or gap-junction proteins (from immunohistochemical staining). STRAP-1 and STRAP-2 are cell-surface tumour antigens. Immunisation with a STRAP protein induces cellular and humoral immune responses against STRAP-expressing cells. STRAP proteins may be used to identify specific antigens, to produce anticancer vaccines and to generate specific antibodies. The antibodies may be used for detection, prognosis, and monitoring of cancers (or susceptibility to cancer), as therapeutic inhibitors or to target therapeutic agents to their site of action. STRAP nucleic acids may be used for recombinant protein production, as diagnostic and prognostic reagents, for identifying STRAP-expressing cells for screening inhibitors of STRAP expression and for therapeutic modulation/inhibition of STRAP expression. Since high levels of STRAP proteins are exposed on the cell surface, they are easily targeted by systemically administered agents, and because they are expressed mainly on prostatic epithelial cells, agents targeted to them should have minimal side effects on other tissues

XX Sequence 34 AA;

Query Match 100.0%; Score 192; DB 3; Length 34;
Best Local Similarity 100.0%; Pred. No. 3.8e-19;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRSVRYKLLNWAYQQVQKQKEDAWIEHDVWRMEI 34
Db 1 RRSVRYKLLNWAYQQVQKQKEDAWIEHDVWRMEI 34

RESULT 2
AAE02787
ID AAE02787 standard; peptide; 34 AA.
XX
AC AAE02787;
XX
DT 06-AUG-2001 (first entry)
XX
DE Extracellular loop #2 of human STEAP-1, suitable for cloning into pFc.
XX
KW Human; cytostatic; antiproliferative; vaccine; gene therapy;
KW six transmembrane epithelial antigen of the prostate-1; STEAP-1;
KW chromosome 7p22.3; cancer; prostate; colon; bladder; pancreatic; ovarian;
KW lung; extracellular loop; serpentine transmembrane antigen.
XX
OS Homo sapiens.
XX
FN WO200140276-A2.
XX
PD 07-JUN-2001.
XX
PF 06-DEC-2000; 2000WO-US033040.
XX
PR 06-DEC-1999; 99US-00455486.
XX
PA (UROG-) UROGENESYS INC.
XX
PI Afar DEH, Hubert RS, Raitano AB, Saffran DC, Mitchell SC;
XX Faris M, Jakobovits A;

XX WPI; 2001-367804/38.
DR
XX
PT New STEAP (six transmembrane epithelial antigen of the prostate)
PT proteins, expressed in human cancers, useful for detecting and treating
PT cancer.
XX
PS Example 19; Page 102; 187pp; English.
XX
CC The present invention relates to human six transmembrane epithelial
CC antigen of the prostate (STEAP) protein. STEAP is a member of cell
CC surface serpentine transmembrane antigens. STEAP gene is used in gene
CC therapy. Inhibiting the development or progression of a cancer (eg.
CC prostate, colon, bladder, lung, ovarian and pancreatic) expressing STEAP
CC or inhibiting growth or killing cells expressing STEAP in a patient,
CC comprises administering a vaccine composition to the patient. Treating a
CC patient with a cancer that expresses STEAP, or inhibiting growth or
CC killing cells expressing STEAP, comprises administering to the patient a
CC vector encoding single chain monoclonal antibody that comprises the
CC variable domains of the heavy and light chains of the monoclonal antibody
CC that specifically binds to STEAP, such that the vector delivers the
CC single chain monoclonal antibody coding sequence to the cancer cells and
CC the encoded single chain monoclonal antibody is expressed
CC intracellularly. The present sequence is extracellular loop of STEAP-1
CC suitable for cloning into pFc, which is used in the invention. STEAP-1
CC gene is located on chromosome 7p22.3
XX
SQ Sequence 34 AA;

Query Match 100.0%; Score 192; DB 4; Length 34;
Best Local Similarity 100.0%; Pred. No. 3.8e-19;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRSVRYKLLNWAYQQVQKQKEDAWIEHDVWRMEI 34
Db 1 RRSVRYKLLNWAYQQVQKQKEDAWIEHDVWRMEI 34

RESULT 3
AAB75315
ID AAB75315 standard; protein; 129 AA.
XX
AC AAB75315;
XX
DT 03-APR-2001 (first entry)
XX
DE Human secreted protein sequence encoded by gene 20 SEQ ID NO:134.
XX
KW Human; immunosuppressive; antiarthritic; antirheumatic; nootropic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW neuroprotective; antibacterial; virucide; fungicide; opthalmological;
KW vulnery; autoimmune disease; hyperproliferative disorder; cancer;
KW cardiovascular disorder; cerebrovascular disorder; infection;
KW nervous system disorder; ocular disorder; chemotaxis; food additive;
KW secreted protein.
XX
OS Homo sapiens.
XX
FN WO200077021-A1.
XX
PD 21-DEC-2000.
XX
PF 01-JUN-2000; 2000WO-US015135.
XX
PR 11-JUN-1999; 99US-0138632P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
PI Rosen CA, Ruben SM, Komatsoulis GA;
XX
XX WPI; 2001-071257/08.
DR N-PSDB; AAF63808.

XX Nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT Parkinson's diseases and cancers.
XX
PS Disclosure; Page 53-54; 530pp; English.
XX
XX This invention relates to polynucleotide sequences AAF63789 - AAF63836
CC which encode human secreted proteins AAB75260 - AAB75287. Included in the
CC invention are protein sequences AAB75288 - AAB75341 which are fragments
CC of the secreted proteins and amino acid sequences with which these
CC fragments share homology. Examples of the activities of the proteins and
CC polynucleotides and the activities of their agonists and antagonists
CC include, immunosuppressive; antithrombotic; antirheumatic;
CC antiproliferative; cytostatic; cardiant; vasotropic; fungicide;
CC neurotropic; neuroprotective; antibacterial; virucide; fungicide;
CC ophthalmological; and vulnary activity. The protein and polynucleotide
CC sequences, their agonists and antagonists may be useful for treating,
CC preventing and diagnosing diseases and disorders such as autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
CC The polypeptides can also be used to aid wound healing and epithelial
CC cell proliferation, to prevent skin aging due to sunburn, to maintain
CC organs before transplantation, for supporting cell culture of primary
CC tissues, to regenerate tissues and in chemotaxis. The polypeptides can
CC also be used as a food additive or preservative to increase or decrease
CC storage capabilities. Included in the invention are oligonucleotides
CC AAF63780 - AAF63788 and peptide AAB75239 which are used in the
CC identification and characterisation of the DNA and protein sequences of
CC the invention
XX
SQ Sequence 129 AA;
Query Match 100.0%; Score 192; DB 4; Length 129;
Best Local Similarity 100.0%; Pred. No. 1.8e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34
Db 60 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 93
RESULT 4
AAB75314
ID AAB75314 standard; protein; 129 AA.
XX AAB75314;
AC
XX
DT 03-APR-2001 (first entry)
XX
DE Gene 20 human secreted protein homologous amino acid sequence #133.
XX Human; immunosuppressive; antithrombotic; antirheumatic; neurotropic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
KW vulnary; autoimmune disease; hyperproliferative disorder; cancer;
KW cardiovascular disorder; cerebrovascular disorder; infection;
KW nervous system disorder; ocular disorder; chemotaxis; food additive;
KW secreted protein.
XX Homo sapiens.
OS
XX WO200077021-A1.
PN
XX
PD 21-DEC-2000.
XX
XX 01-JUN-2000; 2000WO-US015135.
PF
XX 11-JUN-1999; 99US-0138632P.
PR
XX

PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
PI Rosen CA, Ruben SM, Komatsoulis GA;
XX WPI; 2001-071257/08.
DR
XX Nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT Parkinson's diseases and cancers.
XX
PS Disclosure; Page 53-54; 530pp; English.
XX
XX This invention relates to polynucleotide sequences AAF63789 - AAF63836
CC which encode human secreted proteins AAB75260 - AAB75287. Included in the
CC invention are protein sequences AAB75288 - AAB75341 which are fragments
CC of the secreted proteins and amino acid sequences with which these
CC fragments share homology. Examples of the activities of the proteins and
CC polynucleotides and the activities of their agonists and antagonists
CC include, immunosuppressive; antithrombotic; antirheumatic;
CC antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
CC neurotropic; neuroprotective; antibacterial; virucide; fungicide;
CC ophthalmological; and vulnary activity. The protein and polynucleotide
CC sequences, their agonists and antagonists may be useful for treating,
CC preventing and diagnosing diseases and disorders such as autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
CC The polypeptides can also be used to aid wound healing and epithelial
CC cell proliferation, to prevent skin aging due to sunburn, to maintain
CC organs before transplantation, for supporting cell culture of primary
CC tissues, to regenerate tissues and in chemotaxis. The polypeptides can
CC also be used as a food additive or preservative to increase or decrease
CC storage capabilities. Included in the invention are oligonucleotides
CC AAF63780 - AAF63788 and peptide AAB75239 which are used in the
CC identification and characterisation of the DNA and protein sequences of
CC the invention
XX
SQ Sequence 129 AA;
Query Match 100.0%; Score 192; DB 4; Length 129;
Best Local Similarity 100.0%; Pred. No. 1.8e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34
Db 60 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 93
RESULT 5
ABU98426
ID ABU98426 standard; protein; 254 AA.
XX ABU98426;
AC
XX
DT 31-JUL-2003 (first entry)
XX
DE STEAP-1 variant 8P1D4 v.2 #2.
XX
KW STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
KW vaccine.
XX
OS Homo sapiens.
XX
PN WO2003022995-A2.
XX
PD 20-MAR-2003.
XX
PF 06-SEP-2002; 2002WO-US028371.
XX

PR 06-SEP-2001; 2001US-0317840P.
PR 05-APR-2002; 2002US-0370387P.
XX (AGEN-) AGENSYS INC.
XX Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
XX WPI; 2003-313240/30.
XX
XX New composition comprising a substance that modulates the status of a
PT STEAP-1-related protein, useful for treating and detecting cancer.
XX
XX Example 53; Page 169-170; 248pp; English.
XX
XX The invention describes a composition comprising a substance that
CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
CC any of the 15 sequences of 259 amino acids, given in the specification,
CC or a molecule that is modulated by the protein, where the status of the
CC cell that expresses the protein is modulated. The compositions, proteins,
CC polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
XX
SQ Sequence 254 AA;
Query Match 100.0%; Score 192; DB 6; Length 254;
Best Local Similarity 100.0%; Pred. No. 3.8e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RRSYRYKLLNWAYQQVQNKEDAWIEHDVWRMEI 34
Db 185 RRSYRYKLLNWAYQQVQNKEDAWIEHDVWRMEI 218
RESULT 6
ABU98425
ID ABU98425 standard; protein; 254 AA.
XX
AC ABU98425;
XX
XX 31-JUL-2003 (first entry)
XX
XX STEAP-1 variant 8P1D4 v.1 #1.
XX
XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
KW vaccine.
XX
XX Homo sapiens.
OS
XX WO2003022995-A2.
XX
XX 20-MAR-2003.
XX
XX 06-SEP-2002; 2002WO-US028371.
XX
XX 06-SEP-2001; 2001US-0317840P.
PR 05-APR-2002; 2002US-0370387P.
XX
XX (AGEN-) AGENSYS INC.
XX
XX Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
XX WPI; 2003-313240/30.
XX
XX New composition comprising a substance that modulates the status of a
PT STEAP-1-related protein, useful for treating and detecting cancer.
XX
XX Example 53; Page 169-170; 248pp; English.

XX
CC The invention describes a composition comprising a substance that
CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
CC any of the 15 sequences of 259 amino acids, given in the specification,
CC or a molecule that is modulated by the protein, where the status of the
CC cell that expresses the protein is modulated. The compositions, proteins,
CC polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
XX
SQ Sequence 254 AA;
Query Match 100.0%; Score 192; DB 6; Length 254;
Best Local Similarity 100.0%; Pred. No. 3.8e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RRSYRYKLLNWAYQQVQNKEDAWIEHDVWRMEI 34
Db 185 RRSYRYKLLNWAYQQVQNKEDAWIEHDVWRMEI 218
RESULT 7
ABU98429
ID ABU98429 standard; protein; 255 AA.
XX
AC ABU98429;
XX
XX 31-JUL-2003 (first entry)
XX
XX STEAP-1 variant 8P1D4 v.3 #2.
XX
XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
KW vaccine.
XX
XX Homo sapiens.
OS
XX WO2003022995-A2.
XX
XX 20-MAR-2003.
XX
XX 06-SEP-2002; 2002WO-US028371.
XX
XX 06-SEP-2001; 2001US-0317840P.
PR 05-APR-2002; 2002US-0370387P.
XX
XX (AGEN-) AGENSYS INC.
XX
XX Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
XX WPI; 2003-313240/30.
XX
XX New composition comprising a substance that modulates the status of a
PT STEAP-1-related protein, useful for treating and detecting cancer.
XX
XX Example 53; Page 172; 248pp; English.
XX
XX The invention describes a composition comprising a substance that
CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
CC any of the 15 sequences of 259 amino acids, given in the specification,
CC or a molecule that is modulated by the protein, where the status of the
CC cell that expresses the protein is modulated. The compositions, proteins,
CC polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
XX

```

; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-685-166A-879

Query Match      100.0%; Score 86; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WKMKPRNLEDDYL 15
Db      14 WKMKPRNLEDDYL 28

RESULT 3
US-09-134-000C-6624
; Sequence 6624, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6624
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6624

Query Match      53.5%; Score 46; DB 4; Length 216;
Best Local Similarity 61.5%; Pred. No. 4;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      3 MKPRNLEDDYL 15
Db      125 IKPRYLEEGDYI 137

RESULT 4
US-09-489-039A-11367
; Sequence 11367, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747

```

```

; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11367
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11367

Query Match      50.0%; Score 43; DB 4; Length 278;
Best Local Similarity 53.8%; Pred. No. 17;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      2 KMKPRNLEDDY 14
Db      66 RLAPRRELDGY 78

RESULT 5
US-09-134-001C-5278
; Sequence 5278, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5278
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5278

Query Match      49.4%; Score 42.5; DB 4; Length 256;
Best Local Similarity 64.3%; Pred. No. 19;
Matches 9; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY      2 KMKPRNLEDDYL 15
Db      234 KVKPRNN-KEDNYL 246

RESULT 6
US-09-485-632B-6
; Sequence 6, Application US/09485632B
; Patent No. 6605280
; GENERAL INFORMATION:
; APPLICANT: No. 6605280ick, Daniela
; APPLICANT: Dinarello, Charles
; APPLICANT: Rubinstein, Menachem
; APPLICANT: Kim, Soo Hyun
; TITLE OF INVENTION: Interleukin-18 Binding Proteins, their Preparation and
; TITLE OF INVENTION: Use
; FILE REFERENCE: 20993-001
; CURRENT APPLICATION NUMBER: US/09/485,632B
; CURRENT FILING DATE: 2000-10-12
; PRIOR APPLICATION NUMBER: IL98/00379
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: 125463
; PRIOR FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: 122134
; PRIOR FILING DATE: 1997-11-06
; PRIOR APPLICATION NUMBER: 121869
; PRIOR FILING DATE: 1997-09-29
; PRIOR APPLICATION NUMBER: 121639
; PRIOR FILING DATE: 1997-08-27
; PRIOR APPLICATION NUMBER: 121554

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:04:08 ; Search time 9.24419 Seconds
(without alignments)
83.770 Million cell updates/sec

Title: US-10-010-667A-2_COPY_14_28
Perfect score: 86
Sequence: 1 WKMKPRNLEDDYL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PTUS COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86	100.0	339	4	US-09-323-873A-2
2	86	100.0	339	4	US-09-685-166A-879
3	46	53.5	216	4	US-09-134-000C-6624
4	43	50.0	278	4	US-09-489-039A-11367
5	42.5	49.4	256	4	US-09-134-001C-5278
6	40	46.5	197	4	US-09-485-632B-6
7	40	46.5	221	3	US-09-247-373B-54
8	40	46.5	573	4	US-09-252-991A-31334
9	38	44.2	103	4	US-09-621-976-5168
10	38	44.2	193	4	US-09-582-379-4
11	38	44.2	345	1	US-08-403-866-2
12	38	44.2	464	4	US-09-252-991A-29449
13	38	44.2	540	4	US-09-134-000C-4765
14	38	44.2	555	4	US-09-107-532A-6422
15	38	44.2	968	3	US-08-651-939A-7
16	38	44.2	968	3	US-09-385-752-7
17	38	44.2	3969	3	US-08-061-376-5
18	37	43.0	25	1	US-08-468-709B-16
19	37	43.0	25	2	US-08-241-664B-16
20	37	43.0	72	4	US-09-543-681A-7461
21	37	43.0	79	4	US-09-621-976-6283
22	37	43.0	220	4	US-09-489-039A-13425
23	37	43.0	323	2	US-08-969-106-4
24	37	43.0	323	4	US-09-338-125-4
25	37	43.0	416	1	US-08-252-995D-2
26	37	43.0	416	2	US-08-834-108-2
27	37	43.0	464	1	US-08-252-995D-6

```

28      37      43.0      464      2      US-08-834-108-6      Sequence 6, Appli
29      37      43.0      522      4      US-09-446-301A-49      Sequence 49, Appl
30      37      43.0      522      4      US-09-099-932-49      Sequence 49, Appl
31      37      43.0      577      1      US-08-484-105-24      Sequence 24, Appl
32      37      43.0      577      1      US-08-484-106-24      Sequence 24, Appl
33      37      43.0      577      2      US-08-756-317-13      Sequence 13, Appl
34      37      43.0      605      4      US-09-394-645-2       Sequence 2, Appli
35      37      43.0      605      4      US-09-243-560B-2      Sequence 2, Appli
36      37      43.0      711      3      US-08-961-083-82      Sequence 82, Appl
37      37      43.0      711      4      US-09-536-784-82      Sequence 82, Appl
38      37      43.0      751      4      US-09-402-929-2       Sequence 2, Appli
39      37      43.0      752      4      US-09-402-929-6       Sequence 6, Appli
40      37      43.0      925      1      US-08-252-995D-4       Sequence 4, Appli
41      37      43.0      925      2      US-08-834-108-4       Sequence 4, Appli
42      37      43.0      1228     4      US-09-439-313-537     Sequence 537, App
43      37      43.0      1228     4      US-09-636-215-537     Sequence 537, App
44      37      43.0      1228     4      US-09-685-166A-537     Sequence 537, App
45      36      41.9      130      4      US-09-489-039A-7711     Sequence 7711, Ap

```

ALIGNMENTS

```

RESULT 1
US-09-323-873A-2
; Sequence 2, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Kahan Leong
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 129.16USU2
; CURRENT APPLICATION NUMBER: US/09/323,873A
; CURRENT FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-323-873A-2

```

```

Query Match      100.0%; Score 86; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy      1 WKMKPRNLEDDYL 15
Db      14 WKMKPRNLEDDYL 28

```

```

RESULT 2
US-09-685-166A-879
; Sequence 879, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.

```


Db |||||||
 14 WKKPFRNLEDDYL 28

Search completed: July 12, 2004, 14:27:48
Job time : 26.5116 secs

```
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-010-667A-2

Query Match      100.0%; Score 86; DB 14; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WKMKPRRNLEDDYL 15
Db      14 WKMKPRRNLEDDYL 28

RESULT 13
US-10-205-823-397
; Sequence 397, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Monsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; PRIOR FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 397
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-397

Query Match      100.0%; Score 86; DB 14; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WKMKPRRNLEDDYL 15
Db      14 WKMKPRRNLEDDYL 28

RESULT 14
US-10-144-678A-879
; Sequence 879, Application US/10144678A
; Publication No. US20030157089A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Hepler, William T.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals Y de Bassols, Carlota
; APPLICANT: Foy, Teresa M.
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Deng, Ta
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C28
; CURRENT APPLICATION NUMBER: US/10/144,678A
; CURRENT FILING DATE: 2002-08-12
; NUMBER OF SEQ ID NOS: 1033
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-678A-879

Query Match      100.0%; Score 86; DB 14; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WKMKPRRNLEDDYL 15
Db      14 WKMKPRRNLEDDYL 28

RESULT 15
US-10-294-025-879
; Sequence 879, Application US/10294025
; Publication No. US20030185830A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C29
; CURRENT APPLICATION NUMBER: US/10/294,025
; CURRENT FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 1038
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-294-025-879

Query Match      100.0%; Score 86; DB 14; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WKMKPRRNLEDDYL 15
```

```
RESULT 9
US-10-408-009-2
; Sequence 2, Application US/10408009
; Publication No. US20040072196A1
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Stephen C. Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 511582001603
; CURRENT APPLICATION NUMBER: US/10/408,009
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 09/455,486
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/323,873
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-009-2

Query Match      100.0%; Score 86; DB 12; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 WKKPKRRNLEDDYL 15
Db      14 WKKPKRRNLEDDYL 28

RESULT 10
US-10-012-896-879
; Sequence 879, Application US/10012896
; Publication No. US2002018255A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879

Query Match      100.0%; Score 86; DB 12; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 WKKPKRRNLEDDYL 15
Db      14 WKKPKRRNLEDDYL 28

RESULT 11
US-10-011-095-2
; Sequence 2, Application US/10011095
; Publication No. US20030045682A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Hubert, Rene S.
; APPLICANT: Leong, Kahan
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Saffran, Douglas C.
; APPLICANT: Mitchell, Steve Chappell
; TITLE OF INVENTION: ANTIBODIES IMMUNOSPECIFIC FOR STEAP1 (AS AMENDED)
; FILE REFERENCE: 511582001610
; CURRENT APPLICATION NUMBER: US/10/011,095
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA
US-10-011-095-2

Query Match      100.0%; Score 86; DB 14; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 WKKPKRRNLEDDYL 15
Db      14 WKKPKRRNLEDDYL 28

RESULT 12
US-10-010-667A-2
; Sequence 2, Application US/10010667A
; Publication No. US20030055217A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Hubert, Rene S.
; APPLICANT: Leong, Kahan
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Saffran, Douglas C.
; APPLICANT: Mitchell, Steve Chappell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 511582001601
; CURRENT APPLICATION NUMBER: US/10/010,667A
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020187472A1 96572948
US-09-802-520-11
```

```
Query Match      100.0%; Score 86; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 WKMKPRRNLEDDYL 15
        |||||
Db      14 WKMKPRRNLEDDYL 28
```

RESULT 6

```
US-09-895-793-879
; Sequence 879, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolck, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
```

```
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C2
```

```
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
```

```
; SEQ ID NO 879
; LENGTH: 339
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
```

US-09-895-793-879

```
Query Match      100.0%; Score 86; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 WKMKPRRNLEDDYL 15
        |||||
Db      14 WKMKPRRNLEDDYL 28
```

RESULT 7

```
US-09-895-814-879
; Sequence 879, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
```

```
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolck, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
```

```
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C26
```

```
; CURRENT APPLICATION NUMBER: US/09/895,814
```

```
; CURRENT FILING DATE: 2001-06-29
```

```
; NUMBER OF SEQ ID NOS: 990
```

```
; SOFTWARE: FastSEQ for Windows Version 3.0
```

```
; SEQ ID NO 879
```

```
; LENGTH: 339
```

```
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
```

US-09-895-814-879

```
Query Match      100.0%; Score 86; DB 9; Length 339;
```

```
Best Local Similarity 100.0%; Pred. No. 2e-05;
```

```
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 WKMKPRRNLEDDYL 15
        |||||
Db      14 WKMKPRRNLEDDYL 28
```

RESULT 8

US-10-205-267-13

```
; Sequence 13, Application US/10205267
```

```
; Publication No. US20030064397A1
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Spaccake, Kimberly M.
```

```
; APPLICANT: Rickert Paula K.
```

```
; APPLICANT: Lal, Preeti G.
```

```
; APPLICANT: Ison, Craig H.
```

```
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN PROSTATE AND
```

```
; FILE REFERENCE: PV-0008 CIP
```

```
; CURRENT APPLICATION NUMBER: US/10/205,267
```

```
; CURRENT FILING DATE: 2002-07-24
```

```
; NUMBER OF SEQ ID NOS: 13
```

```
; SOFTWARE: PEEL Program
```

```
; SEQ ID NO 13
```

```
; LENGTH: 339
```

```
; TYPE: PRT
```

```
; ORGANISM:
```

```
; PUBLICATION INFORMATION:
```

```
; DATABASE ACCESSION NUMBER: Genbank ID No: 96572948
```

US-10-205-267-13

```
Query Match      100.0%; Score 86; DB 12; Length 339;
```

```
Best Local Similarity 100.0%; Pred. No. 2e-05;
```

```
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 WKMKPRRNLEDDYL 15
        |||||
Db      14 WKMKPRRNLEDDYL 28
```

```
Query Match      100.0%; Score 86; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKKKPRRNLEDDYL 15
Db 1 WKKKPRRNLEDDYL 15

RESULT 2
US-09-759-143-879
; Sequence 879, Application US/09759143
; Patent No. US2002002248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John H.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-143-879

Query Match      100.0%; Score 86; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKKKPRRNLEDDYL 15
Db 14 WKKKPRRNLEDDYL 28

RESULT 3
US-09-780-669-879
; Sequence 879, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John H.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun

Query Match      100.0%; Score 86; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKKKPRRNLEDDYL 15
Db 14 WKKKPRRNLEDDYL 28

RESULT 4
US-09-822-827-879
; Sequence 879, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-822-827-879

Query Match      100.0%; Score 86; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKKKPRRNLEDDYL 15
Db 14 WKKKPRRNLEDDYL 28

RESULT 5
US-09-802-520-11
; Sequence 11, Application US/09802520
; Publication No. US20020187472A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Faris, Mary
; APPLICANT: Chen, Huel-Mei
; APPLICANT: Ison, Craig H.
; TITLE OF INVENTION: STEAP-RELATED PROTEIN
; FILE REFERENCE: PC-0037 US
; CURRENT APPLICATION NUMBER: US/09/802,520
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PERL Program
; SEQ ID NO 11
; LENGTH: 339
; TYPE: PRT
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:10:33 ; Search time 26.5116 Seconds
(without alignments)
176.480 Million cell updates/sec

Title: US-10-010-667a-2_COPY_14_28

Perfect score: 86
Sequence: 1 WKKFRRNLEEDYL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1279676 seqs, 311918243 residues

Total number of hits satisfying chosen parameters: 1279676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86	100.0	15	14	US-10-165-044-22
2	86	100.0	339	9	US-09-759-143-879
3	86	100.0	339	9	US-09-780-669-879
4	86	100.0	339	9	US-09-822-827-879
5	86	100.0	339	9	US-09-802-520-11
6	86	100.0	339	9	US-09-895-793-879
7	86	100.0	339	9	US-09-895-814-879
8	86	100.0	339	12	US-10-205-267-13
9	86	100.0	339	12	US-10-408-009-2
10	86	100.0	339	13	US-10-012-896-879
11	86	100.0	339	14	US-10-011-095-2
12	86	100.0	339	14	US-10-010-667a-2
13	86	100.0	339	14	US-10-205-823-397
14	86	100.0	339	14	US-10-144-678A-879
15	86	100.0	339	14	US-10-294-025-879

16	86	100.0	339	15	US-10-239-607-37	Sequence 37, Appl
17	86	100.0	339	15	US-10-295-027-714	Sequence 714, App
18	86	100.0	339	15	US-10-295-027-1347	Sequence 1347, Ap
19	86	100.0	368	12	US-10-425-114-72779	Sequence 72779, A
20	86	100.0	375	14	US-10-165-044-2	Sequence 2, Appli
21	46	53.5	561	12	US-10-424-599-248099	Sequence 248099,
22	45	52.3	2078	16	US-10-437-963-130585	Sequence 130585,
23	44	51.2	66	12	US-10-424-599-243298	Sequence 243298,
24	44	51.2	193	14	US-10-189-346-60	Sequence 60, Appl
25	43	50.0	137	16	US-10-437-963-128802	Sequence 128802,
26	43	50.0	398	16	US-10-437-963-165124	Sequence 165124,
27	43	50.0	2040	15	US-10-276-968-13	Sequence 13, Appl
28	42	48.8	411	16	US-10-437-963-204350	Sequence 204350,
29	42	48.8	442	15	US-10-369-493-17721	Sequence 17721, A
30	42	48.8	1018	9	US-09-801-574-32	Sequence 32, Appl
31	42	48.8	1019	9	US-09-801-574-76	Sequence 76, Appl
32	42	48.8	1019	14	US-10-132-861-2	Sequence 2, Appli
33	42	48.8	1137	12	US-10-425-114-63876	Sequence 63876, A
34	41.5	48.3	208	15	US-10-158-034-67	Sequence 67, Appl
35	41.5	48.3	305	12	US-10-425-114-43994	Sequence 43994, A
36	41.5	48.3	528	12	US-10-424-599-259292	Sequence 259292,
37	41.5	48.3	705	12	US-10-425-114-43788	Sequence 43788, A
38	41	47.7	32	12	US-10-671-403-203	Sequence 203, App
39	41	47.7	32	12	US-10-671-419-203	Sequence 203, App
40	41	47.7	32	12	US-10-670-844-203	Sequence 203, App
41	41	47.7	32	12	US-10-671-134-203	Sequence 203, App
42	41	47.7	32	12	US-10-673-098-203	Sequence 203, App
43	41	47.7	32	16	US-10-672-638-203	Sequence 203, App
44	41	47.7	32	16	US-10-673-127-203	Sequence 203, App
45	41	47.7	32	16	US-10-670-817-203	Sequence 203, App

ALIGNMENTS

RESULT 1

US-10-165-044-22
; Sequence 22, Application US/10165044
; Publication No. US20030149531A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas Saffran
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Steven Chappell Mitchell
; APPLICANT: Mary Faris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 51158-20016.02
; CURRENT APPLICATION NUMBER: US/10165,044
; CURRENT FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: US 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: US 60/091,183
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 09/455,486
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: WO 99/62941
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/33040
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-165-044-22

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:01:53 ; Search time 7.5 seconds
(without alignments)
192.383 Million cell updates/sec

Title: US-10-010-667A-2_COPY_14_28
Perfect score: 86
Sequence: 1 WKMKPRNLEDDYL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45.5	52.9	237	2 T50202	conserved hypothe
2	45	52.3	375	1 NRECD	ribonuclease III (
3	45	52.3	375	2 A99943	RNase D, processes
4	45	52.3	375	2 E85791	RNase D, processes
5	44	51.2	532	2 T02539	hypotheical prote
6	43	50.0	312	2 T15371	hypotheical prote
7	43	50.0	539	2 T39150	probable heat shoc
8	42	48.8	124	2 AH1754	bacteriophage prot
9	42	48.8	384	2 T49020	hypotheical prote
10	42	48.8	416	2 D96992	hypotheical prote
11	42	48.8	442	2 D36718	dihydroliipoamide S
12	42	48.8	451	2 T49021	hypotheical prote
13	42	48.8	847	2 JC4836	alpha-glucuronidas
14	42	48.8	1832	2 AC2594	glutamate synthase
15	42	48.8	1858	2 D97376	hypotheical prote
16	41	47.7	260	2 S75010	ABC-type transport
17	41	47.7	271	2 C84550	hypotheical prote
18	41	47.7	403	2 G75287	NADH oxidase-relat
19	41	47.7	580	2 T21493	hypotheical prote
20	41	47.7	1187	2 T46637	transcription fact
21	40.5	47.1	639	2 F71258	probable heat shoc
22	40.5	47.1	658	2 S76909	hypotheical prote
23	40.5	47.1	1415	1 EDBEGA	immediate-early pr
24	40	46.5	36	2 B31872	retinoic acid-bind
25	40	46.5	138	2 I51265	xCRABP - African c
26	40	46.5	361	2 T49337	hypotheical prote
27	40	46.5	430	2 T23899	hypotheical prote
28	40	46.5	474	2 T00943	hypotheical prote
29	40	46.5	526	2 F85086	hypotheical prote

30	40	46.5	549	1 F69361	arginyl-tRNA synth
31	40	46.5	884	2 H83322	hypotheical prote
32	40	46.5	961	2 T01167	hypotheical prote
33	40	46.5	1268	2 T18955	hypotheical prote
34	39	45.3	90	2 H90665	probable Ogr famil
35	39	45.3	90	2 C85516	probable activator
36	39	45.3	371	2 C64499	isocitrate dehydro
37	39	45.3	426	2 E83981	pyruvate dehydrog
38	39	45.3	426	2 B95519	protein T286.2 lfm
39	39	45.3	449	2 G89841	hypotheical prote
40	39	45.3	479	2 B69764	transcription regu
41	39	45.3	512	2 T23035	hypotheical prote
42	39	45.3	652	2 B84568	probable calmoduli
43	39	45.3	713	2 JCS870	poly(beta-D-mannur
44	39	45.3	976	2 A97104	zn-dependent metal
45	39	45.3	1019	2 T00117	dve protein - frul

ALIGNMENTS

RESULT 1

T50202
conserved hypotheical protein SPAC25B8.15c [imported] - fission yeast (Schizosaccharomy
C;Species: Schizosaccharomyces pombe
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jun-2000
C;Accession: T50202
R;Barrell, B.G.; Rajandream, M.A.; McDougall, R.C.; McLean, J.; Harris, D.
submitted to the EMBL Data Library, November 1999
A;Reference number: Z25045
A;Accession: T50202
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-237 <BAR>
A;Cross-references: EMBL:AL133225; PIDN:CAB61781.1; GSPDB:GN00066; SPDB:SPAC25B8.15c
A;Experimental source: strain 972h(-); cosmid c25B8
C;Genetics:
A;Gene: SPDB:SPAC25B8.15c
A;Map position: 1

Query Match 52.9%; Score 45.5; DB 2; Length 237;
Best Local Similarity 66.7%; Pred.No. 4.3;
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 2 WKMKPRNLEE-DDYL 15
Db 220 WKMKPRKLRNMDYYL 234
:|||||
:|||||

RESULT 2

NRECD
ribonuclease III (EC 3.1.26.3) rnd - Escherichia coli (strain K-12)
N;Alternate names: ribonuclease D
C;Species: Escherichia coli
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 01-Mar-2002
C;Accession: S01223; A30431; S41590; D64941; S42849; D21915
R;Zhang, J.; Deutscher, M.P.
Nucleic Acids Res. 16, 6265-6278, 1988
A;Title: Escherichia coli RNase D: sequencing of the rnd structural gene and purification
A;Reference number: S01223; MUID:88289400; PMID:3041371
A;Accession: S01223
A;Molecule type: DNA
A;Residues: 1-375 <ZHA>
A;Cross-references: EMBL:X07055; NID:g42770; PIDN:CAA30098.1; PID:g581215
A;Accession: A30431
A;Molecule type: protein
A;Residues: 1-6 <ZH2>
R;Fulda, M.; Heinz, E.; Wolter, F.P.
Mol. Gen. Genet. 242, 241-249, 1994
A;Title: The fadD gene of Escherichia coli K12 is located close to rnd at 39.6 min of t
A;Reference number: S41590; MUID:94150456; PMID:8107670
A;Accession: S41590
A;Molecule type: DNA

```
A;Residues: 1-38 <FUL>
A;Cross-references: GB:X70994; NID:g433478; PIDN:CAA50322.1; PID:g581071
A;Experimental source: strain K12
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shaoh, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: D64941
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-375 <BLAT>
A;Cross-references: GB:AE000274; GB:U00096; NID:G1788089; PIDN:AACT4874.1; PID:g1788105;
A;Experimental source: strain K-12, substrain MG1655
R;Fulda, M.
submitted to the EMBL Data Library, February 1993
A;Reference number: S42848
A;Accession: S42849
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 'L', 2-38 <FU2>
A;Cross-references: EMBL:X70994
C;Genetics:
A;Gene: rnd
A;Map position: 40 min
A;Start codon: TTG
C;Superfamily: ribonuclease D
C;Keywords: exonuclease; hydrolase
F;1-375/Product: ribonuclease D #status experimental <MAT>

Query Match 52.3%; Score 45; DB 1; Length 375;
Best Local Similarity 63.6%; Pred. No. 8.5;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WKMKPRRNLEE 11
|||:||||
Db 342 WKLKPNQLPE 352

RESULT 3
A99943
RNase D, processes tRNA precursor [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C;Accession: A99943
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: A99943
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-375 <HAY>
A;Cross-references: GB:BA000007; PIDN:BA035936.1; PID:g13361980; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RMD 0509952
C;Genetics:
A;Gene: EC82513
C;Superfamily: ribonuclease D

Query Match 52.3%; Score 45; DB 2; Length 375;
Best Local Similarity 63.6%; Pred. No. 8.5;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WKMKPRRNLEE 11
|||:||||
Db 342 WKLKPNQLPE 352

RESULT 4
B85791
RNase D, processes tRNA precursor [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
```

```
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: E85791
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, L.;
Iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: E85791
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-375 <STO>
A;Cross-references: GB:AE005174; NID:g12515844; PIDN:AA056793.1; GSPDB:GN00145; UWGP:Z
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: rnd
C;Superfamily: ribonuclease D

Query Match 52.3%; Score 45; DB 2; Length 375;
Best Local Similarity 63.6%; Pred. No. 8.5;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WKMKPRRNLEE 11
|||:||||
Db 342 WKLKPNQLPE 352

RESULT 5
T02539
Hypothetical protein At2g37730 [imported] - Arabidopsis thaliana
N;Alternate names: hypothetical protein F13M22.23
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
C;Accession: T02539; D84796
R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,
submitted to the EMBL Data Library, June 1998
A;Description: Arabidopsis thaliana chromosome II BAC F13M22 genomic sequence.
A;Reference number: Z14677
A;Accession: T02539
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-532 <ROU>
A;Cross-references: EMBL:AC004684; NID:g3236234; PID:g3236255
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: D84796
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-532 <STO>
A;Cross-references: GB:AE002093; NID:g3236255; PIDN:AA023643.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g37730; F13M22.23
A;Map position: 2
A;Introns: 165/2; 295/3; 478/3
C;Superfamily: Arabidopsis hypothetical protein F13M22.23

Query Match 51.2%; Score 44; DB 2; Length 532;
Best Local Similarity 55.6%; Pred. No. 18;
Matches 10; Conservative 0; Mismatches 2; Indels 6; Gaps 1;

Qy 1 WKMKPRR-----NLEED 12
|||:||||
Db 476 WKMAPRQCCEIVNSEED 493

RESULT 6
T15371
Hypothetical protein C01F1.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
```


C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T15371
R;Johnson, D.
submitted to the EMBL Data Library, May 1996
A;Description: The sequence of C. elegans cosmid C01F1.
A;Reference number: Z18338
C;Accession: T15371
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-312 <OJH>
A;Cross-references: EMBL:U58761; NID:gl330391; PID:gl330396; PIDN:AAB00716.1; GSPDB:GN00178
A;Experimental source: strain Bristol N2; clone C01F1
C;Genetics:
A;Gene: CESP:C01F1.2
A;Map position: 2
A;Introns: 36/2; 271/1

Query Match 50.0%; Score 43; DB 2; Length 312;
Best Local Similarity 63.6%; Pred. No. 15;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 PRNLEDDYL 15
|||:||||:
Db 250 PRNLEDDYL 260

RESULT 7
T39150
probable heat shock transcription factor - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T39150
R;Oliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A;Reference number: Z21748
A;Accession: T39150
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-539 <OLJ>
A;Cross-references: EMBL:Z99168; PIDN:CAB16301.1; GSPDB:GN00066; SPDB:SPAC8C9.14
A;Experimental source: strain 972H-; cosmid c8C9
C;Genetics:
A;Gene: SPDB:SPAC8C9.14
A;Map position: 1
A;Introns: 10/3; 40/3; 67/2; 86/3

Query Match 50.0%; Score 43; DB 2; Length 539;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 WKMKPRRNLEDDYL 15
|||:||||:
Db 364 WKRPRLIVDEDEL 378

RESULT 8
AH1754
bacteriophage protein homolog lin2581 [imported] - Listeria innocua (strain Clp11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AH1754
R;Glaser, P.; Frangul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative Genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AH1754
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-124 <GLA>

A;Cross-references: GB:AL592022; PIDN:CAC97808.1; PID:gl6415103; GSPDB:GN00178
A;Experimental source: strain Clp11262
C;Genetics:
A;Gene: lin2581

Query Match 48.8%; Score 42; DB 2; Length 124;
Best Local Similarity 46.7%; Pred. No. 8.6;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 WKMKPRRNLEDDYL 15
|||:||||:
Db 15 WKRRKKILRRDEYL 29

RESULT 9
T49020
hypothetical protein F3C22.70 - Arabidopsis thaliana (fragment)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C;Accession: T49020
R;Purnelle, B.; Masuy, D.; Goffeau, A.; Boutry, M.; Mewes, H.W.; Rudd, S.; Lemcke, K.; P
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z25013
A;Accession: T49020
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-384 <PUR>
A;Cross-references: EMBL:ALJ53912; GSPDB:GN00061; ATSP:F3C22.70
A;Experimental source: cultivar Columbia; BAC clone F3C22
C;Genetics:
A;Gene: ATSP:F3C22.70
A;Map position: 3
A;Introns: 61/1; 252/3; 281/3

Query Match 48.8%; Score 42; DB 2; Length 384;
Best Local Similarity 42.9%; Pred. No. 28;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 WKMKPRRNLEDDY 14
|||:||||:
Db 46 WKLVNLEFDSDDY 59

RESULT 10
D96692
hypothetical protein T12I7.9 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: D96692
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D96692
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-416 <STO>
A;Cross-references: GB:AE005173; NID:gl1054573; PIDN:AAG27848.1; GSPDB:GN00141
C;Genetics:
A;Gene: T12I7.9
A;Map position: 1

Query Match 48.8%; Score 42; DB 2; Length 416;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 WKMKPRNLEEDDY 14
||| :|||
Db 40 WKYVPRDLDEADF 53

RESULT 11

D36718

dihydrolipoamide S-acetyltransferase (EC 2.3.1.12) precursor - Bacillus subtilis
N;Alternate names: pyruvate dehydrogenase complex, E2 component
C;Species: Bacillus subtilis

C;Date: 12-Apr-1991 #sequence_revision 12-Apr-1991 #text_change 20-Jun-2000
C;Accession: D36718, B69674

R;Hemilae: H.; Palva, A.; Paulin, L.; Arvidson, S.; Palva, I.
J. Bacteriol. 172, 5052-5063, 1990

A;Title: Secretory S complex of Bacillus subtilis: sequence analysis and identity to pyv
A;Reference number: A36718; MUID:90368558; PMID:1697575

A;Accession: D36718

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-442 <HEM>

A;Cross-references: GB:M57435; GB:M31542; NID:g143375; PIDN:AAA62683.1; PID:g143379

R;Kunit, P.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Etrington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gall

isch, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue

Y. M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,

A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron

teuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Togononi, A.; Tosato, V.; Uchiyama

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A;Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: B69674

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-442 <KUN>

A;Cross-references: GB:Z99111; GB:AL009126; NID:g2633699; PIDN:CAB13333.1; PID:g2633831

A;Experimental source: strain 168

C;Genetics:

A;Gene: pdhC

C;Superfamily: dihydrolipoamide acetyltransferase; lipoyl/biotin-binding homology

C;Keywords: acyltransferase; coenzyme A; oxidoreductase

F;4-77/Domain: lipoyl/biotin-binding homology <LBP>

F;413,417/Active site: His, Asp #status predicted

Query Match 48.8%; Score 42; DB 2; Length 442;
Best Local Similarity 46.7%; Pred. No. 33;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 WKMKPRNLEEDDY 15
||| :|||
Db 22 WYVPRNDEDDVL 36

RESULT 12

T49021

hypothetical protein F3C22.80 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000

C;Accession: T49021

R;Purnelle, B.; Masny, D.; Goffeau, A.; Boutry, M.; Mewes, H.W.; Rudd, S.; Lemcke, K.; M

submitted to the Protein Sequence Database, April 2000

A;Reference number: Z25013

A;Accession: T49021

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-451 <PUR>

A;Cross-references: EMBL:AL353912; GSPDB:GN00061; ATSP:F3C22.80

A;Experimental source: cultivar Columbia; BAC clone F3C22

C;Genetics:

A;Gene: ATSP:F3C22.80

A;Map position: 3

A;Introns: 4/2; 294/3; 339/3

Query Match 48.8%; Score 42; DB 2; Length 451;
Best Local Similarity 42.9%; Pred. No. 33;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 WKMKPRNLEEDDY 14
||| :|||
Db 52 WKLVNLEPDSDDY 55

RESULT 13

JC4836

alpha-glucuronidase (EC 3.2.1.-) precursor - fungus (Trichoderma reesei)

N;Alternate names: GLRI

C;Species: Trichoderma reesei

C;Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 03-Nov-2003

C;Accession: JC4836

R;Margolles-Clark, E.; Saloheimo, M.; Slika-aho, M.; Penttilae, M.

Gene 172, 171-172, 1996

A;Title: The alpha-glucuronidase-encoding gene of Trichoderma reesei.

A;Reference number: JC4836; MUID:96257277; PMID:8654984

A;Accession: JC4836

A;Molecule type: mRNA

A;Residues: 1-847 <MAR>

A;Cross-references: EMBL:Z68706; NID:g1419337; PID:e218512; PID:g1419338

C;Comment: This enzyme releases glucuronic acid attached to xylose units of xylan.

C;Genetics:

A;Gene: glr1

C;Superfamily: alpha-glucuronidase

C;Keywords: glycosidase; hydrolase

F;1-19/Domain: signal sequence #status predicted <SIG>

F;20-847/Product: alpha-glucuronidase #status predicted <MAP>

Query Match 48.8%; Score 42; DB 2; Length 847;
Best Local Similarity 53.8%; Pred. No. 65;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 KMKPRNLEEDDY 14
||| :|||
Db 108 KLSPLNLEEDDY 120

RESULT 14

AC2594

glutamate synthase large subunit gltB [imported] - Agrobacterium tumefaciens (strain C58)

C;Species: Agrobacterium tumefaciens

C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

C;Accession: AC2594

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A;Reference number: AB2577; MUID:21608550; PMID:11743193

A;Accession: AC2594

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1832 <KUR>

A;Cross-references: GB:AE008688; PIDN:AAL41169.1; PID:g17738468; GSPDB:GN00186

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: gltB

A;Map position: circular chromosome

Query Match 48.8%; Score 42; DB 2; Length 1832;
Best Local Similarity 53.8%; Pred. No. 1.4e+02;

This Page Blank (uspto)

this Page Blank (uspto)

; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; FILE REFERENCE: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14765
; LENGTH: 996
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-14765

Query Match 48.8%; Score 42; DB 6; Length 996;
Best Local Similarity 53.8%; Pred. No. 96;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 WKMKPRNLEEDD 13
Db 622 WKLVKSNLEED 634

Search completed: July 12, 2004, 14:25:04
Job time : 10.4186 secs

US-10-812-849-25

RESULT 15
US-10-170-205E-14765
: Sequence 14765, Application US/10170205E

; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18326
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-18326

Query Match 100.0%; Score 86; DB 6; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKMKPRRNLEDDYL 15
Db 24 WKMKPRRNLEDDYL 38
|||||

RESULT 7

US-10-857-785-35
; Sequence 35, Application US/10857785
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Stephen C. Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 511582001607
; CURRENT APPLICATION NUMBER: US/10/857,785
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: US 09/455,486
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 09/323,873
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: US 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: US 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-857-785-35

Query Match 87.8%; Score 75.5; DB 6; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 WKMKP-RRNLEDDYL 15
Db 1 WKMKPRRNLEDDYL 16
|||||

RESULT 8

US-10-856-109-35
; Sequence 35, Application US/10856109
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Stephen C. Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 511582001606
; CURRENT APPLICATION NUMBER: US/10/856,109
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: US 09/455,486

; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: US 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-856-109-35

Query Match 87.8%; Score 75.5; DB 6; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 WKMKP-RRNLEDDYL 15
Db 1 WKMKPRRNLEDDYL 16
|||||

RESULT 9

US-10-752-421-35
; Sequence 35, Application US/10752421
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Stephen C. Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 511582001612
; CURRENT APPLICATION NUMBER: US/10/752,421
; CURRENT FILING DATE: 2004-01-05
; PRIOR APPLICATION NUMBER: US 09/455,486
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: US 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-752-421-35

Query Match 87.8%; Score 75.5; DB 6; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 WKMKP-RRNLEDDYL 15
Db 1 WKMKPRRNLEDDYL 16
|||||

RESULT 10

US-10-724-972A-5260
; Sequence 5260, Application US/10724972A
; GENERAL INFORMATION:
; APPLICANT: Doucette-Stamm, Lynn
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH03-16
; CURRENT APPLICATION NUMBER: US/10/724,972A

```

; PRIOR APPLICATION NUMBER: PCT/US01/09410
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,929
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-239-607A-37

Query Match
Best Local Similarity 100.0%; Score 86; DB 6; Length 339;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKMKPRNLEDDYL 15
Db 14 WKMKPRNLEDDYL 28

RESULT 3
US-10-857-785-2
; Sequence 2, Application US/10857785
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Stephen C. Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 511582001607
; CURRENT APPLICATION NUMBER: US/10/857,785
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: US 09/455,486
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: US 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-857-785-2

Query Match
Best Local Similarity 100.0%; Score 86; DB 6; Length 339;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKMKPRNLEDDYL 15
Db 14 WKMKPRNLEDDYL 28

RESULT 4
US-10-856-109-2
; Sequence 2, Application US/10856109
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Stephen C. Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 511582001606
; CURRENT APPLICATION NUMBER: US/10/856,109

```

```

; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: US 09/455,486
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: US 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-856-109-2

Query Match
Best Local Similarity 100.0%; Score 86; DB 6; Length 339;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKMKPRNLEDDYL 15
Db 14 WKMKPRNLEDDYL 28

RESULT 5
US-10-752-421-2
; Sequence 2, Application US/10752421
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Stephen C. Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 511582001612
; CURRENT APPLICATION NUMBER: US/10/752,421
; CURRENT FILING DATE: 2004-01-05
; PRIOR APPLICATION NUMBER: US 09/455,486
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: US 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-752-421-2

Query Match
Best Local Similarity 100.0%; Score 86; DB 6; Length 339;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKMKPRNLEDDYL 15
Db 14 WKMKPRNLEDDYL 28

RESULT 6
US-10-170-205E-18326
; Sequence 18326, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:08:08 ; Search time 9.4186 Seconds
(without alignments)
118.046 Million cell updates/sec

Title: US-10-010-667a-2_COPY_14_28

Perfect score: 86

Sequence: 1 WKKPRNLEDDYL 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 393080 seqs, 74121830 residues

Total number of hits satisfying chosen parameters: 393080

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- Pending Patents AA New:*
- 1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
 - 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
 - 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
 - 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
 - 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
 - 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
 - 7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	86	100.0	339	6	US-10-821-801-741
2	86	100.0	339	6	US-10-239-607A-37
3	86	100.0	339	6	US-10-857-785-2
4	86	100.0	339	6	US-10-856-109-2
5	86	100.0	339	6	US-10-752-421-2
6	86	100.0	349	6	US-10-170-205E-18326
7	75.5	87.8	16	6	US-10-857-785-35
8	75.5	87.8	16	6	US-10-856-109-35
9	75.5	87.8	16	6	US-10-752-421-35
10	42.5	49.4	256	6	US-10-724-972A-5260
11	42	48.8	331	1	PCT-US04-19153-25
12	42	48.8	331	6	US-10-812-849-25
13	42	48.8	817	6	US-10-767-701-4432
14	42	48.8	897	6	US-10-170-205E-14827
15	42	48.8	996	6	US-10-170-205E-14765
16	42	48.8	1832	7	US-60-581-351-5639
17	42	48.8	1858	7	US-60-581-351-5641
18	41	47.7	32	5	US-09-716-964B-203
19	41	47.7	32	6	US-10-671-106-203
20	41	47.7	327	6	US-10-170-205E-23733
21	40	46.5	72	5	US-09-248-796A-21475
22	40	46.5	197	5	US-09-790-338B-6
23	40	46.5	234	6	US-10-170-205E-16541
24	40	46.5	317	6	US-10-767-471-744
25	40	46.5	317	6	US-10-767-471-748
26	40	46.5	317	6	US-10-170-205E-28535

27	40	46.5	360	6	US-10-767-471-747	Sequence 747, App
28	40	46.5	360	6	US-10-170-205E-27412	Sequence 27412, A
29	40	46.5	365	7	US-60-556-841-5722	Sequence 5722, Ap
30	40	46.5	465	6	US-10-767-471-745	Sequence 745, App
31	40	46.5	770	6	US-10-767-471-750	Sequence 750, App
32	40	46.5	770	6	US-10-821-801-848	Sequence 848, App
33	40	46.5	1012	6	US-10-767-471-749	Sequence 749, App
34	40	46.5	1012	6	US-10-170-205E-27413	Sequence 27413, A
35	39.5	45.9	490	1	PCT-US03-24982A-125	Sequence 125, App
36	39.5	45.9	687	7	US-60-551-161-260	Sequence 260, App
37	39	45.3	142	7	US-60-565-632-11886	Sequence 11886, A
38	39	45.3	142	7	US-60-579-062-11886	Sequence 11886, A
39	39	45.3	321	6	US-10-170-205E-24904	Sequence 24904, A
40	39	45.3	349	7	US-60-546-745-23	Sequence 23, Appl
41	39	45.3	485	1	PCT-US04-19924-32	Sequence 32, Appl
42	39	45.3	485	6	US-10-873-467-32	Sequence 32, Appl
43	39	45.3	485	7	US-60-551-121-100	Sequence 100, App
44	39	45.3	485	7	US-60-581-351-13766	Sequence 13766, A
45	39	45.3	528	1	PCT-US02-22858A-357	Sequence 357, App

ALIGNMENTS

RESULT 1

US-10-821-801-741
; Sequence 741, Application US/10821801
; GENERAL INFORMATION:
; APPLICANT: Hinzmann, Bernd
; APPLICANT: Rosenthal, Andre
; APPLICANT: Hermann, Klaus
; APPLICANT: Heiden, Esmeralda
; APPLICANT: Pilarsky, Christian
; APPLICANT: Brummendorf, Thomas
; APPLICANT: Staub, Eike
; APPLICANT: Ropcke, Stefan
; APPLICANT: Mennerich, Detlev
; APPLICANT: Kinnemann, Henrik
; APPLICANT: Li, Xinzhong
; TITLE OF INVENTION: Human nucleic acid sequences from lung tumours
; FILE REFERENCE: 00154/002001
; CURRENT APPLICATION NUMBER: US/10/821,801
; CURRENT FILING DATE: 2004-04-09
; PRIOR APPLICATION NUMBER: DE 103 16 701.3
; PRIOR FILING DATE: 2003-04-09
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 741
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-821-801-741

Query Match 100.0%; Score 86; DB 6; Length 339;
Best Local Similarity 100.0%; Pred.No. 1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WKKPRNLEDDYL 15
| | | | | | | | | | | | | | | | | |
Db 14 WKKPRNLEDDYL 28

RESULT 2

US-10-239-607A-37
; Sequence 37, Application US/10239607A
; GENERAL INFORMATION:
; APPLICANT: Saatcioglu, Fahri
; TITLE OF INVENTION: Novel Prostate-Specific or
; TITLE OF INVENTION: Testis-Specific Nucleic Acid Molecules, Polypeptides, and
; TITLE OF INVENTION: Diagnostic and Therapeutic Methods
; FILE REFERENCE: 50218/003002
; CURRENT APPLICATION NUMBER: US/10/239,607A
; CURRENT FILING DATE: 2003-03-20

This Page Blank (uspto)

This Page Blank (uspto)

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 MKPRNLEDDYL 15
::||:||||
Db 356 LRPLRTVETDDYL 368

RESULT 15

D97376
Hypothetical protein AGR_C_235 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C:Accession: D97376
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorllo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: D97376
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1858 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK85965.1; PID:g15155024; GSPDB:GM00169
C:Genetics:
A:Gene: AGR_C_235
A:Map position: circular chromosome

Query Match 48.8%; Score 42; DB 2; Length 1858;

Best Local Similarity 53.8%; Pred. No. 1.5e+02;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 MKPRNLEDDYL 15
::||:||||
Db 382 LRPLRTVETDDYL 394

Search completed: July 12, 2004, 14:11:17
Job time : 8.5 secs

This Page Blank (uspto)

This Page Blank (uspto)

Search completed: July 12, 2004, 14:08:02
Job time : 5.7093 secs

```
KW Complete proteome.
FT SITE 113 123 "HIGH" REGION.
SQ SEQUENCE 549 AA; 62861 MW; 39DF41CC0B9AD210 CRC64;

Query Match 46.5%; Score 40; DB 1; Length 549;
Best Local Similarity 46.2%; Pred. No. 46;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 WKMKPRRLNEEDD 13
Db 377 WKILSERDMEDE 389

RESULT 15
GPT1_CANAL STANDARD; PRT; 553 AA.
AC 074248;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative polyamine transporter.
GN GPT1.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21179897; PubMed=11284011;
RA McNemar M.D.; Gorman J.A.; Buckley H.R.;
RT "Isolation of a gene encoding a putative polyamine transporter from
RT Candida albicans, GPT1."
RL Yeast 18:555-561(2001).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: Belongs to the amino acid permease family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF080132; AAC31569.1; -.
DR InterPro; IPR002293; AA/rel_permease1.
DR InterPro; IPR004840; AAC_permease.
DR InterPro; IPR004841; Permease_region.
DR Pfam; PF00324; aa_permeases; 1.
DR PROSITE; PS00218; AMINO_ACID_PERMEASE_1; FALSE_NEG.
KW Transport; Transmembrane.
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 68 88 POTENTIAL.
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 163 183 POTENTIAL.
FT TRANSMEM 190 210 POTENTIAL.
FT TRANSMEM 240 260 POTENTIAL.
FT TRANSMEM 280 300 POTENTIAL.
FT TRANSMEM 333 353 POTENTIAL.
FT TRANSMEM 397 417 POTENTIAL.
FT TRANSMEM 453 473 POTENTIAL.
FT TRANSMEM 485 505 POTENTIAL.
SQ SEQUENCE 553 AA; 60633 MW; 197DACB76C5FBC9E CRC64;

Query Match 46.5%; Score 40; DB 1; Length 553;
Best Local Similarity 46.2%; Pred. No. 46;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 WKMKPRRLNEEDD 13
Db 508 WYVGPRTNISEED 520
```

RT "Interleukin-18 binding protein: a novel modulator of the Th1 cytokine
response.";
RL Immunity 10:127-136(1999).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM A), SEQUENCE OF 29-33; 37-43; 46-52;
55-58; 61-72; 89-93; 105-142 AND 149-167, AND TISSUE SPECIFICITY.
RX MEDLINE=99192308; PubMed=10094485;
RA Aizawa Y., Akita K., Tanitai M., Torigoe K., Mori T., Nishida Y.,
Uchida S., Nukada Y., Tanimoto T., Ikegami H., Ikeda M., Kurimoto M.;
RT "Cloning and expression of interleukin-18 binding protein.";
RL FEBS Lett. 445:338-342(1999).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM A).
RX MEDLINE=99263157; PubMed=10329540;
RA Xiang Y., Moss B.;
RT Identification of human and mouse homologs of the MC51L-53L-54L
family of secreted glycoproteins encoded by the Mollusca contagiosum
poxvirus.";
RL Virology 257:297-302(1999).
CC -!- FUNCTION: Binds to IL-18 and inhibits its activity. Functions as
an inhibitor of the early Th1 cytokine response.
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=C; Synonyms=IL-18BPC;
CC IsoId=O95998-1; Sequence=Displayed;
CC Name=A; Synonyms=IL-18BPA;
CC IsoId=O95998-2; Sequence=VSP_002514;
CC Name=B; Synonyms=IL-18BPB;
CC IsoId=O95998-3; Sequence=VSP_002515, VSP_002516;
CC -!- TISSUE SPECIFICITY: Strongly expressed in heart, lung, placenta
and spleen.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF110798; AAD17187.1; -
DR EMBL; AF110798; AAD17188.1; -
DR EMBL; AF110798; AAD17189.1; -
DR EMBL; AF110799; AAD17190.1; -
DR EMBL; AF110800; AAD17191.1; -
DR EMBL; AF110801; AAD17192.1; -
DR EMBL; AF110802; AAD17193.1; -
DR EMBL; AF110803; AAD17194.1; -
DR EMBL; AF110804; AAD17195.1; -
DR EMBL; AF110805; AAD17196.1; -
DR EMBL; AF110806; AAD17197.1; -
DR EMBL; AF110807; AAD17198.1; -
DR EMBL; AF110808; AAD17199.1; -
DR EMBL; AF110809; AAD17200.1; -
DR EMBL; AF110810; AAD17201.1; -
DR EMBL; AF110811; AAD17202.1; -
DR EMBL; AF110812; AAD17203.1; -
DR EMBL; AF110813; AAD17204.1; -
DR EMBL; AF110814; AAD17205.1; -
DR EMBL; AF110815; AAD17206.1; -
DR EMBL; AF110816; AAD17207.1; -
DR EMBL; AF110817; AAD17208.1; -
DR EMBL; AF110818; AAD17209.1; -
DR EMBL; AF110819; AAD17210.1; -
DR EMBL; AF110820; AAD17211.1; -
DR EMBL; AF110821; AAD17212.1; -
DR EMBL; AF110822; AAD17213.1; -
DR EMBL; AF110823; AAD17214.1; -
DR EMBL; AF110824; AAD17215.1; -
DR EMBL; AF110825; AAD17216.1; -
DR EMBL; AF110826; AAD17217.1; -
DR EMBL; AF110827; AAD17218.1; -
DR EMBL; AF110828; AAD17219.1; -
DR EMBL; AF110829; AAD17220.1; -
DR EMBL; AF110830; AAD17221.1; -
DR EMBL; AF110831; AAD17222.1; -
DR EMBL; AF110832; AAD17223.1; -
DR EMBL; AF110833; AAD17224.1; -
DR EMBL; AF110834; AAD17225.1; -
DR EMBL; AF110835; AAD17226.1; -
DR EMBL; AF110836; AAD17227.1; -
DR EMBL; AF110837; AAD17228.1; -
DR EMBL; AF110838; AAD17229.1; -
DR EMBL; AF110839; AAD17230.1; -
DR EMBL; AF110840; AAD17231.1; -
DR EMBL; AF110841; AAD17232.1; -
DR EMBL; AF110842; AAD17233.1; -
DR EMBL; AF110843; AAD17234.1; -
DR EMBL; AF110844; AAD17235.1; -
DR EMBL; AF110845; AAD17236.1; -
DR EMBL; AF110846; AAD17237.1; -
DR EMBL; AF110847; AAD17238.1; -
DR EMBL; AF110848; AAD17239.1; -
DR EMBL; AF110849; AAD17240.1; -
DR EMBL; AF110850; AAD17241.1; -
DR EMBL; AF110851; AAD17242.1; -
DR EMBL; AF110852; AAD17243.1; -
DR EMBL; AF110853; AAD17244.1; -
DR EMBL; AF110854; AAD17245.1; -
DR EMBL; AF110855; AAD17246.1; -
DR EMBL; AF110856; AAD17247.1; -
DR EMBL; AF110857; AAD17248.1; -
DR EMBL; AF110858; AAD17249.1; -
DR EMBL; AF110859; AAD17250.1; -
DR EMBL; AF110860; AAD17251.1; -
DR EMBL; AF110861; AAD17252.1; -
DR EMBL; AF110862; AAD17253.1; -
DR EMBL; AF110863; AAD17254.1; -
DR EMBL; AF110864; AAD17255.1; -
DR EMBL; AF110865; AAD17256.1; -
DR EMBL; AF110866; AAD17257.1; -
DR EMBL; AF110867; AAD17258.1; -
DR EMBL; AF110868; AAD17259.1; -
DR EMBL; AF110869; AAD17260.1; -
DR EMBL; AF110870; AAD17261.1; -
DR EMBL; AF110871; AAD17262.1; -
DR EMBL; AF110872; AAD17263.1; -
DR EMBL; AF110873; AAD17264.1; -
DR EMBL; AF110874; AAD17265.1; -
DR EMBL; AF110875; AAD17266.1; -
DR EMBL; AF110876; AAD17267.1; -
DR EMBL; AF110877; AAD17268.1; -
DR EMBL; AF110878; AAD17269.1; -
DR EMBL; AF110879; AAD17270.1; -
DR EMBL; AF110880; AAD17271.1; -
DR EMBL; AF110881; AAD17272.1; -
DR EMBL; AF110882; AAD17273.1; -
DR EMBL; AF110883; AAD17274.1; -
DR EMBL; AF110884; AAD17275.1; -
DR EMBL; AF110885; AAD17276.1; -
DR EMBL; AF110886; AAD17277.1; -
DR EMBL; AF110887; AAD17278.1; -
DR EMBL; AF110888; AAD17279.1; -
DR EMBL; AF110889; AAD17280.1; -
DR EMBL; AF110890; AAD17281.1; -
DR EMBL; AF110891; AAD17282.1; -
DR EMBL; AF110892; AAD17283.1; -
DR EMBL; AF110893; AAD17284.1; -
DR EMBL; AF110894; AAD17285.1; -
DR EMBL; AF110895; AAD17286.1; -
DR EMBL; AF110896; AAD17287.1; -
DR EMBL; AF110897; AAD17288.1; -
DR EMBL; AF110898; AAD17289.1; -
DR EMBL; AF110899; AAD17290.1; -
DR EMBL; AF110900; AAD17291.1; -
DR EMBL; AF110901; AAD17292.1; -
DR EMBL; AF110902; AAD17293.1; -
DR EMBL; AF110903; AAD17294.1; -
DR EMBL; AF110904; AAD17295.1; -
DR EMBL; AF110905; AAD17296.1; -
DR EMBL; AF110906; AAD17297.1; -
DR EMBL; AF110907; AAD17298.1; -
DR EMBL; AF110908; AAD17299.1; -
DR EMBL; AF110909; AAD17300.1; -
DR EMBL; AF110910; AAD17301.1; -
DR EMBL; AF110911; AAD17302.1; -
DR EMBL; AF110912; AAD17303.1; -
DR EMBL; AF110913; AAD17304.1; -
DR EMBL; AF110914; AAD17305.1; -
DR EMBL; AF110915; AAD17306.1; -
DR EMBL; AF110916; AAD17307.1; -
DR EMBL; AF110917; AAD17308.1; -
DR EMBL; AF110918; AAD17309.1; -
DR EMBL; AF110919; AAD17310.1; -
DR EMBL; AF110920; AAD17311.1; -
DR EMBL; AF110921; AAD17312.1; -
DR EMBL; AF110922; AAD17313.1; -
DR EMBL; AF110923; AAD17314.1; -
DR EMBL; AF110924; AAD17315.1; -
DR EMBL; AF110925; AAD17316.1; -
DR EMBL; AF110926; AAD17317.1; -
DR EMBL; AF110927; AAD17318.1; -
DR EMBL; AF110928; AAD17319.1; -
DR EMBL; AF110929; AAD17320.1; -
DR EMBL; AF110930; AAD17321.1; -
DR EMBL; AF110931; AAD17322.1; -
DR EMBL; AF110932; AAD17323.1; -
DR EMBL; AF110933; AAD17324.1; -
DR EMBL; AF110934; AAD17325.1; -
DR EMBL; AF110935; AAD17326.1; -
DR EMBL; AF110936; AAD17327.1; -
DR EMBL; AF110937; AAD17328.1; -
DR EMBL; AF110938; AAD17329.1; -
DR EMBL; AF110939; AAD17330.1; -
DR EMBL; AF110940; AAD17331.1; -
DR EMBL; AF110941; AAD17332.1; -
DR EMBL; AF110942; AAD17333.1; -
DR EMBL; AF110943; AAD17334.1; -
DR EMBL; AF110944; AAD17335.1; -
DR EMBL; AF110945; AAD17336.1; -
DR EMBL; AF110946; AAD17337.1; -
DR EMBL; AF110947; AAD17338.1; -
DR EMBL; AF110948; AAD17339.1; -
DR EMBL; AF110949; AAD17340.1; -
DR EMBL; AF110950; AAD17341.1; -
DR EMBL; AF110951; AAD17342.1; -
DR EMBL; AF110952; AAD17343.1; -
DR EMBL; AF110953; AAD17344.1; -
DR EMBL; AF110954; AAD17345.1; -
DR EMBL; AF110955; AAD17346.1; -
DR EMBL; AF110956; AAD17347.1; -
DR EMBL; AF110957; AAD17348.1; -
DR EMBL; AF110958; AAD17349.1; -
DR EMBL; AF110959; AAD17350.1; -
DR EMBL; AF110960; AAD17351.1; -
DR EMBL; AF110961; AAD17352.1; -
DR EMBL; AF110962; AAD17353.1; -
DR EMBL; AF110963; AAD17354.1; -
DR EMBL; AF110964; AAD17355.1; -
DR EMBL; AF110965; AAD17356.1; -
DR EMBL; AF110966; AAD17357.1; -
DR EMBL; AF110967; AAD17358.1; -
DR EMBL; AF110968; AAD17359.1; -
DR EMBL; AF110969; AAD17360.1; -
DR EMBL; AF110970; AAD17361.1; -
DR EMBL; AF110971; AAD17362.1; -
DR EMBL; AF110972; AAD17363.1; -
DR EMBL; AF110973; AAD17364.1; -
DR EMBL; AF110974; AAD17365.1; -
DR EMBL; AF110975; AAD17366.1; -
DR EMBL; AF110976; AAD17367.1; -
DR EMBL; AF110977; AAD17368.1; -
DR EMBL; AF110978; AAD17369.1; -
DR EMBL; AF110979; AAD17370.1; -
DR EMBL; AF110980; AAD17371.1; -
DR EMBL; AF110981; AAD17372.1; -
DR EMBL; AF110982; AAD17373.1; -
DR EMBL; AF110983; AAD17374.1; -
DR EMBL; AF110984; AAD17375.1; -
DR EMBL; AF110985; AAD17376.1; -
DR EMBL; AF110986; AAD17377.1; -
DR EMBL; AF110987; AAD17378.1; -
DR EMBL; AF110988; AAD17379.1; -
DR EMBL; AF110989; AAD17380.1; -
DR EMBL; AF110990; AAD17381.1; -
DR EMBL; AF110991; AAD17382.1; -
DR EMBL; AF110992; AAD17383.1; -
DR EMBL; AF110993; AAD17384.1; -
DR EMBL; AF110994; AAD17385.1; -
DR EMBL; AF110995; AAD17386.1; -
DR EMBL; AF110996; AAD17387.1; -
DR EMBL; AF110997; AAD17388.1; -
DR EMBL; AF110998; AAD17389.1; -
DR EMBL; AF110999; AAD17390.1; -
DR EMBL; AF111000; AAD17391.1; -
DR EMBL; AF111001; AAD17392.1; -
DR EMBL; AF111002; AAD17393.1; -
DR EMBL; AF111003; AAD17394.1; -
DR EMBL; AF111004; AAD17395.1; -
DR EMBL; AF111005; AAD17396.1; -
DR EMBL; AF111006; AAD17397.1; -
DR EMBL; AF111007; AAD17398.1; -
DR EMBL; AF111008; AAD17399.1; -
DR EMBL; AF111009; AAD17400.1; -
DR EMBL; AF111010; AAD17401.1; -
DR EMBL; AF111011; AAD17402.1; -
DR EMBL; AF111012; AAD17403.1; -
DR EMBL; AF111013; AAD17404.1; -
DR EMBL; AF111014; AAD17405.1; -
DR EMBL; AF111015; AAD17406.1; -
DR EMBL; AF111016; AAD17407.1; -
DR EMBL; AF111017; AAD17408.1; -
DR EMBL; AF111018; AAD17409.1; -
DR EMBL; AF111019; AAD17410.1; -
DR EMBL; AF111020; AAD17411.1; -
DR EMBL; AF111021; AAD17412.1; -
DR EMBL; AF111022; AAD17413.1; -
DR EMBL; AF111023; AAD17414.1; -
DR EMBL; AF111024; AAD17415.1; -
DR EMBL; AF111025; AAD17416.1; -
DR EMBL; AF111026; AAD17417.1; -
DR EMBL; AF111027; AAD17418.1; -
DR EMBL; AF111028; AAD17419.1; -
DR EMBL; AF111029; AAD17420.1; -
DR EMBL; AF111030; AAD17421.1; -
DR EMBL; AF111031; AAD17422.1; -
DR EMBL; AF111032; AAD17423.1; -
DR EMBL; AF111033; AAD17424.1; -
DR EMBL; AF111034; AAD17425.1; -
DR EMBL; AF111035; AAD17426.1; -
DR EMBL; AF111036; AAD17427.1; -
DR EMBL; AF111037; AAD17428.1; -
DR EMBL; AF111038; AAD17429.1; -
DR EMBL; AF111039; AAD17430.1; -
DR EMBL; AF111040; AAD17431.1; -
DR EMBL; AF111041; AAD17432.1; -
DR EMBL; AF111042; AAD17433.1; -
DR EMBL; AF111043; AAD17434.1; -
DR EMBL; AF111044; AAD17435.1; -
DR EMBL; AF111045; AAD17436.1; -
DR EMBL; AF111046; AAD17437.1; -
DR EMBL; AF111047; AAD17438.1; -
DR EMBL; AF111048; AAD17439.1; -
DR EMBL; AF111049; AAD17440.1; -
DR EMBL; AF111050; AAD17441.1; -
DR EMBL; AF111051; AAD17442.1; -
DR EMBL; AF111052; AAD17443.1; -
DR EMBL; AF111053; AAD17444.1; -
DR EMBL; AF111054; AAD17445.1; -
DR EMBL; AF111055; AAD17446.1; -
DR EMBL; AF111056; AAD17447.1; -
DR EMBL; AF111057; AAD17448.1; -
DR EMBL; AF111058; AAD17449.1; -
DR EMBL; AF111059; AAD17450.1; -
DR EMBL; AF111060; AAD17451.1; -
DR EMBL; AF111061; AAD17452.1; -
DR EMBL; AF111062; AAD17453.1; -
DR EMBL; AF111063; AAD17454.1; -
DR EMBL; AF111064; AAD17455.1; -
DR EMBL; AF111065; AAD17456.1; -
DR EMBL; AF111066; AAD17457.1; -
DR EMBL; AF111067; AAD17458.1; -
DR EMBL; AF111068; AAD17459.1; -
DR EMBL; AF111069; AAD17460.1; -
DR EMBL; AF111070; AAD17461.1; -
DR EMBL; AF111071; AAD17462.1; -
DR EMBL; AF111072; AAD17463.1; -
DR EMBL; AF111073; AAD17464.1; -
DR EMBL; AF111074; AAD17465.1; -
DR EMBL; AF111075; AAD17466.1; -
DR EMBL; AF111076; AAD17467.1; -
DR EMBL; AF111077; AAD17468.1; -
DR EMBL; AF111078; AAD17469.1; -
DR EMBL; AF111079; AAD17470.1; -
DR EMBL; AF111080; AAD17471.1; -
DR EMBL; AF111081; AAD17472.1; -
DR EMBL; AF111082; AAD17473.1; -
DR EMBL; AF111083; AAD17474.1; -
DR EMBL; AF111084; AAD17475.1; -
DR EMBL; AF111085; AAD17476.1; -
DR EMBL; AF111086; AAD17477.1; -
DR EMBL; AF111087; AAD17478.1; -
DR EMBL; AF111088; AAD17479.1; -
DR EMBL; AF111089; AAD17480.1; -
DR EMBL; AF111090; AAD17481.1; -
DR EMBL; AF111091; AAD17482.1; -
DR EMBL; AF111092; AAD17483.1; -
DR EMBL; AF111093; AAD17484.1; -
DR EMBL; AF111094; AAD17485.1; -
DR EMBL; AF111095; AAD17486.1; -
DR EMBL; AF111096; AAD17487.1; -
DR EMBL; AF111097; AAD17488.1; -
DR EMBL; AF111098; AAD17489.1; -
DR EMBL; AF111099; AAD17490.1; -
DR EMBL; AF111100; AAD17491.1; -
DR EMBL; AF111101; AAD17492.1; -
DR EMBL; AF111102; AAD17493.1; -
DR EMBL; AF111103; AAD17494.1; -
DR EMBL; AF111104; AAD17495.1; -
DR EMBL; AF111105; AAD17496.1; -
DR EMBL; AF111106; AAD17497.1; -
DR EMBL; AF111107; AAD17498.1; -
DR EMBL; AF111108; AAD17499.1; -
DR EMBL; AF111109; AAD17500.1; -
DR EMBL; AF111110; AAD17501.1; -
DR EMBL; AF111111; AAD17502.1; -
DR EMBL; AF111112; AAD17503.1; -
DR EMBL; AF111113; AAD17504.1; -
DR EMBL; AF111114; AAD17505.1; -
DR EMBL; AF111115; AAD17506.1; -
DR EMBL; AF111116; AAD17507.1; -
DR EMBL; AF111117; AAD17508.1; -
DR EMBL; AF111118; AAD17509.1; -
DR EMBL; AF111119; AAD17510.1; -
DR EMBL; AF111120; AAD17511.1; -
DR EMBL; AF111121; AAD17512.1; -
DR EMBL; AF111122; AAD17513.1; -
DR EMBL; AF111123; AAD17514.1; -
DR EMBL; AF111124; AAD17515.1; -
DR EMBL; AF111125; AAD17516.1; -
DR EMBL; AF111126; AAD17517.1; -
DR EMBL; AF111127; AAD17518.1; -
DR EMBL; AF111128; AAD17519.1; -
DR EMBL; AF111129; AAD17520.1; -
DR EMBL; AF111130; AAD17521.1; -
DR EMBL; AF111131; AAD17522.1; -
DR EMBL; AF111132; AAD17523.1; -
DR EMBL; AF111133; AAD17524.1; -
DR EMBL; AF111134; AAD17525.1; -
DR EMBL; AF111135; AAD17526.1; -
DR EMBL; AF111136; AAD17527.1; -
DR EMBL; AF111137; AAD17528.1; -
DR EMBL; AF111138; AAD17529.1; -
DR EMBL; AF111139; AAD17530.1; -
DR EMBL; AF111140; AAD17531.1; -
DR EMBL; AF111141; AAD17532.1; -
DR EMBL; AF111142; AAD17533.1; -
DR EMBL; AF111143; AAD17534.1; -
DR EMBL; AF111144; AAD17535.1; -
DR EMBL; AF111145; AAD17536.1; -
DR EMBL; AF111146; AAD17537.1; -
DR EMBL; AF111147; AAD17538.1; -
DR EMBL; AF111148; AAD17539.1; -
DR EMBL; AF111149; AAD17540.1; -
DR EMBL; AF111150; AAD17541.1; -
DR EMBL; AF111151; AAD17542.1; -
DR EMBL; AF111152; AAD17543.1; -
DR EMBL; AF111153; AAD17544.1; -
DR EMBL; AF111154; AAD17545.1; -
DR EMBL; AF111155; AAD17546.1; -
DR EMBL; AF111156; AAD17547.1; -
DR EMBL; AF111157; AAD17548.1; -
DR EMBL; AF111158; AAD17549.1; -
DR EMBL; AF111159; AAD17550.1; -
DR EMBL; AF111160; AAD17551.1; -
DR EMBL; AF111161; AAD17552.1; -
DR EMBL; AF111162; AAD17553.1; -
DR EMBL; AF111163; AAD17554.1; -
DR EMBL; AF111164; AAD17555.1; -
DR EMBL; AF111165; AAD17556.1; -
DR EMBL; AF111166; AAD17557.1; -
DR EMBL; AF111167; AAD17558.1; -
DR EMBL; AF111168; AAD17559.1; -
DR EMBL; AF111169; AAD17560.1; -
DR EMBL; AF111170; AAD17561.1; -
DR EMBL; AF111171; AAD17562.1; -
DR EMBL; AF111172; AAD17563.1; -
DR EMBL; AF111173; AAD17564.1; -
DR EMBL; AF111174; AAD17565.1; -
DR EMBL; AF111175; AAD17566.1; -
DR EMBL; AF111176; AAD17567.1; -
DR EMBL; AF111177; AAD17568.1; -
DR EMBL; AF111178; AAD17569.1; -
DR EMBL; AF111179; AAD17570.1; -
DR EMBL; AF111180; AAD17571.1; -
DR EMBL; AF111181; AAD17572.1; -
DR EMBL; AF111182; AAD17573.1; -
DR EMBL; AF111183; AAD17574.1; -
DR EMBL; AF111184; AAD17575.1; -
DR EMBL; AF111185; AAD17576.1; -
DR EMBL; AF111186; AAD17577.1; -
DR EMBL; AF111187; AAD17578.1; -
DR EMBL; AF111188; AAD17579.1; -
DR EMBL; AF111189; AAD17580.1; -
DR EMBL; AF111190; AAD17581.1; -
DR EMBL; AF111191; AAD17582.1; -
DR EMBL; AF111192; AAD17583.1; -
DR EMBL; AF111193; AAD17584.1; -
DR EMBL; AF111194; AAD17585.1; -
DR EMBL; AF111195; AAD17586.1; -
DR EMBL; AF111196; AAD17587.1; -
DR EMBL; AF111197; AAD17588.1; -
DR EMBL; AF111198; AAD17589.1; -
DR EMBL; AF111199; AAD17590.1; -
DR EMBL; AF111200; AAD17591.1; -
DR EMBL; AF111201; AAD17592.1; -
DR EMBL; AF111202; AAD17593.1; -
DR EMBL; AF111203; AAD17594.1; -
DR EMBL; AF111204; AAD17595.1; -
DR EMBL; AF111205; AAD17596.1; -
DR EMBL; AF111206; AAD17597.1; -
DR EMBL; AF111207; AAD17598.1; -
DR EMBL; AF111208; AAD17599.1; -
DR EMBL; AF111209; AAD17600.1; -
DR EMBL; AF111210; AAD17601.1; -
DR EMBL; AF111211; AAD17602.1; -
DR EMBL; AF111212; AAD17603.1; -
DR EMBL; AF111213; AAD17604.1; -
DR EMBL; AF111214; AAD17605.1; -
DR EMBL; AF111215; AAD17606.1; -
DR EMBL; AF111216; AAD17607.1; -
DR EMBL; AF111217; AAD17608.1; -

```

DR EMBL; M75729; AAA46111.1; -.
DR InterPro; IPR005205; Herpes_ICP4_C.
DR InterPro; IPR005206; Herpes_ICP4_N.
DR Pfam; PF03585; Herpes_ICP4_C; 1.
DR Pfam; PF03584; Herpes_ICP4_N; 1.
KW Early protein; Transcription regulation; Trans-acting factor;
KW Activator; DNA-binding; Phosphorylation; Nuclear protein.
FT DOMAIN 155 200
FT SER/PRO-RICH.
SQ SEQUENCE 1415 AA; 154936 MW; C0846F7BEF4D1126 CRC64;

Query Match 47.1%; Score 40.5; DB 1; Length 1415;
Best Local Similarity 60.8%; Pred. No. 1e+02;
Matches 9; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

OY 2 KMKPR-RNLEEDYL 15
Db 222 KASPRTRKLEDEDYL 236

RESULT 11
ID RET4_CHICK STANDARD; PRT; 36 AA.
AC P30370;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Retinoic acid-binding protein II, cellular (CRABP-II) (Fragment).
GN CRABP2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP MEDLINE=89087453; PubMed=2849937;
RA Kitamoto T., Momoi T., Momoi M.;
RT "The presence of a novel cellular retinoic acid-binding protein in
RT chick embryos: purification and partial characterization.";
RL Biochem. Biophys. Res. Commun. 157:1302-1308(1988).
RN [2]
RP ERRATUM.
RA Kitamoto T., Momoi T., Momoi M.;
RL Biochem. Biophys. Res. Commun. 159:371-371(1989).
CC -!- FUNCTION: CYTOSOLIC CRABPS MAY REGULATE THE ACCESS OF RETINOIC
CC ACID TO THE NUCLEAR RETINOIC ACID RECEPTORS.
CC -!- FUNCTION: CRABP-II MAY PARTICIPATE IN A REGULATORY FEEDBACK
CC MECHANISM TO CONTROL THE ACTION OF RETINOIC ACID ON CELL
CC DIFFERENTIATION.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Embryo.
CC -!- SIMILARITY: Belongs to the fatty-acid binding protein (FABP)
CC family.
DR PIR; B31872; B31872.
DR HSSP; P29373; ICBS.
DR InterPro; IPR000463; Fatty_acid_BP.
DR InterPro; IPR000566; Lipocalin_cytfabp.
DR Pfam; PF00061; lipocalin; 1.
DR PROSITE; PS00214; FABP; 1.
KW Vitamin A; Transport.
FT NON TER 36
FT SEQUENCE 36 AA; 3997 MW; 2BC1C154B2FD9904 CRC64;

Query Match 46.5%; Score 40; DB 1; Length 36;
Best Local Similarity 63.6%; Pred. No. 2.5;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 WKMKPRRNLEE 11
Db 7 WKMKSENFE 17

RESULT 12
ID RET3_XENLA STANDARD; PRT; 137 AA.
AC P50558;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Retinoic acid-binding protein, cellular (CRABP).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE=95034308; PubMed=7947321;
RA Ho L., Mercola M., Gudas L.J.;
RT "Xenopus laevis cellular retinoic acid-binding protein: temporal and
RT spatial expression pattern during early embryogenesis.";
RL Mech. Dev. 47:53-64(1994).
CC -!- FUNCTION: CYTOSOLIC CRABPS MAY REGULATE THE ACCESS OF RETINOIC
CC ACID TO THE NUCLEAR RETINOIC ACID RECEPTORS.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the fatty-acid binding protein (FABP)
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; S74933; AAB32580.1; -.
DR PIR; I51265; I51265.
DR HSSP; P29373; ICBS.
DR InterPro; IPR000463; Fatty acid BP.
DR InterPro; IPR000566; Lipocalin_cytfabp.
DR Pfam; PF00061; lipocalin; 1.
DR PRINTS; PR00178; PATTYACIDBP.
DR PROSITE; PS00214; FABP; 1.
KW Vitamin A; Transport.
FT INIT MET 0
FT SEQUENCE 137 AA; 15524 MW; 2193DDE7B095EC4A CRC64;

Query Match 46.5%; Score 40; DB 1; Length 137;
Best Local Similarity 63.6%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 WKMKPRRNLEE 11
Db 7 WKMKSENFE 17

RESULT 13
ID I18B_HUMAN STANDARD; PRT; 197 AA.
AC O95958; O95993; O96027; Q9UBR7;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Interleukin-18 binding protein precursor (IL-18BP).
GN IL18BP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.; SEQUENCE OF 29-68, FUNCTION, ALTERNATIVE SPLICING,
RP AND TISSUE SPECIFICITY.
RX MEDLINE=99146382; PubMed=10023777;
RA Novick D., Kim S.-H., Fantuzzi G., Reznikov L.L., Dinarello C.A.,
RA Rubinstein M.;

```

```
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser Thr pkin_AS.
DR InterPro; IPR002999; Tudor.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF00867; Tudor; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00333; Tudor; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; FALSE NEG.
DR PROSITE; PS00108; PROTEIN KINASE ST; FALSE NEG.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS03044; Tudor; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Coiled coil.
FT DOMAIN 78 137 TUDOR.
FT DOMAIN 105 242 TNASE-LIKE.
FT DOMAIN 298 355 COILED COIL (POTENTIAL).
FT DOMAIN 710 1019 PROTEIN KINASE.
FT BIND 716 724
FT BINDING 737 737
FT CONFLICT 300 300 I -> F (IN REF. 2).
FT CONFLICT 509 509 F -> Y (IN REF. 2).
FT CONFLICT 621 623 KKI -> NKS (IN REF. 2).
FT CONFLICT 715 715 Y -> C (IN REF. 2).
FT CONFLICT 820 820 V -> A (IN REF. 2).
FT CONFLICT 948 948 K -> I (IN REF. 2).
FT CONFLICT 963 963 A -> G (IN REF. 2).
FT CONFLICT 1000 1000 T -> P (IN REF. 2).
FT CONFLICT 1010 1010 P -> T (IN REF. 2).
SQ SEQUENCE 1019 AA; 115729 MW; 5B58244A98D1590A CRC64;

Query Match 48.8%; Score 42; DB 1; Length 1019;
Best Local Similarity 53.8%; Pred. No. 41;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 WKMKPRNLEEDD 13
DB 645 WKLVKSNLEEDS 657

RESULT 9
HTPG TREPA
ID _HTPG_TREPA STANDARD; PRT; 639 AA.
AC O8349;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chaperone protein htpG (Heat shock protein htpG) (High temperature
DE protein G)
DE HTPG OR TP0984.
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artachon P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RL Science 281:375-388(1998).
CC -!- FUNCTION: Molecular chaperone. Has ATPase activity (By
CC similarity).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the heat shock protein 90 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
-----
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE001265; AAC65938.1; -.
CC PIR; F71258; F71258.
CC HSSP; P02829; 1AH8.
CC TIGR; TP0984; -.
CC HAMAP; MF_00505; -. 1.
CC InterPro; IPR003594; ATPbind_ATPase.
CC InterPro; IPR001404; Hsp90.
CC Pfam; PF02518; HATPase_c; 1.
CC Pfam; PF00183; HSP90; 2.
CC PRINTS; PR00775; HEATSHOCK90.
CC SMART; SM00387; HATPase_c; 1.
CC PROSITE; PS00298; HSP90; 1.
KW Chaperone; ATP-binding; Heat shock; Complete proteome.
FT DOMAIN 1 348 A; SUBSTRATE-BINDING (BY SIMILARITY).
FT DOMAIN 349 565 B (BY SIMILARITY).
FT DOMAIN 566 639 C.
SQ SEQUENCE 639 AA; 72937 MW; 3E8FDBAC2282C31D CRC64;

Query Match 47.1%; Score 40.5; DB 1; Length 639;
Best Local Similarity 50.0%; Pred. No. 45;
Matches 7; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

QY 1 WKMKPRNLEEDDY 14
DB 246 WK-RPKSELKEDDY 258

RESULT 10
ICP4 HSVMG
ID _ICP4_HSVMG STANDARD; PRT; 1415 AA.
AC Q02362;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Trans-acting transcriptional activator protein ICP4 (Immediate-early
DE protein IE175).
DE ICP4.
GN ICP4.
OS Marek's disease herpesvirus (strain GA) (MDHV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10388;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92351564; PubMed=1322594;
RA Anderson A.S., Francesconi A., Morgan R.W.;
RT "Complete nucleotide sequence of the Marek's disease virus ICP4
RT gene."
RL Virology 189:657-667(1992).
CC -!- FUNCTION: THIS IE PROTEIN IS A MULTIFUNCTIONAL PROTEIN CAPABLE
CC OF MIGRATING TO THE NUCLEUS, BINDING TO DNA, TRANS-ACTIVATING
CC OTHER VIRAL GENES, AND AUTO-REGULATING ITS OWN SYNTHESIS. IT IS
CC REQUIRED FOR THE SWITCH FROM IMMEDIATE-EARLY TO EARLY MODE OF
CC GENE EXPRESSION.
CC -!- SUBCELLULAR LOCATION: Nucleus of infected cells.
CC -!- PTM: A long stretch of serine residues may be a major site of
CC phosphorylation.
CC -!- SIMILARITY: BELONGS TO THE ICP4/IE140/IE180 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
```

```
CC -!- SIMILARITY: BELONGS TO FAMILY 67 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z68706; CAA92949.1; -.
CC FIRM; JC4836; JC4836.
CC InterPro; IPR005154; Glyco_hydro_67.
CC Pfam; PF03648; Glyco_hydro_67; 1.
CC Xylan degradation; Hydrolase; Glycosidase; Glycoprotein; Signal.
CC SIGNAL 1 19 POTENTIAL.
CC CHAIN 20 847 ALPHA-GLUCURONIDASE.
CC CARBOHYD 52 52 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 321 321 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 353 353 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 586 586 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 692 692 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 740 740 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 767 767 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SEQUENCE 847 AA; 93424 MW; 214AAE487382FDB4 CRC64;
CC -----
CC Query Match 48.8%; Score 42; DB 1; Length 847;
CC Best Local Similarity 53.8%; Pred.No.34;
CC Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
CC -----
CC QY 2 KMKPRNLEEDY 14
CC |:|:|:|:|
CC Db 108 KLSPLKLNKEDGY 120
CC -----
CC RESULT 7
CC ST31_MOUSE
CC ID ST31_MOUSE STANDARD; PRT; 1018 AA.
CC AC Q99WU1;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DE Serine/threonine protein kinase 31 (EC 2.7.1.37).
CC GN STK31.
CC OS Mus musculus (Mouse).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC OX NCBI_TaxID=10090;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Testis;
CC RX MEDLINE=21175748; PubMed=11279525;
CC RA Wang P.J., McCarrey J.R., Yang F., Page D.C.;
CC RT "An abundance of X-linked genes expressed in spermatogonia.";
CC RL Nat. Genet. 27:422-426(2001).
CC CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC CC -!- TISSUE SPECIFICITY: Testis specific. Expressed only in male germ
CC cells.
CC CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC CC -!- SIMILARITY: Contains 1 Tudor domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF285580; AAK31959.1; -.
CC MGD; MG1:1924735; STK31.
CC InterPro; IPR008191; Maternal_tudor.
CC -----
```

```
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002999; Tudor.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00567; Tudor; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00333; Tudor; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS0304; Tudor; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Coiled coil.
FT DOMAIN 78 137 TUDOR.
FT DOMAIN 105 245 TNASE-LIKE.
FT DOMAIN 298 358 COILED COIL (POTENTIAL).
FT DOMAIN 711 1018 PROTEIN_KINASE.
FT NP_BIND 717 725 ATP (BY SIMILARITY).
FT BINDING 738 738 ATP (BY SIMILARITY).
SQ SEQUENCE 1018 AA; 115027 MW; 1206284ED4E90816 CRC64;
CC -----
CC Query Match 48.8%; Score 42; DB 1; Length 1018;
CC Best Local Similarity 53.8%; Pred.No.41;
CC Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
CC -----
CC QY 1 WKMPPRNLEEDD 13
CC |:|:|:|:|
CC Db 646 WKLVEKSNLEED 658
CC -----
CC RESULT 8
CC ST31_HUMAN
CC ID ST31_HUMAN STANDARD; PRT; 1019 AA.
CC AC Q9BXU1; Q9BXH8;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Serine/threonine protein kinase 31 (EC 2.7.1.37) (Serine/threonine-
CC protein kinase NYD-SPK).
CC GN STK31.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC OX NCBI_TaxID=9606;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Testis;
CC RX MEDLINE=21175748; PubMed=11279525;
CC RA Wang P.J., McCarrey J.R., Yang F., Page D.C.;
CC RT "An abundance of X-linked genes expressed in spermatogonia.";
CC RL Nat. Genet. 27:422-426(2001).
CC RN [2]
CC RP SEQUENCE OF 24-1019 FROM N.A.
CC RC TISSUE=Testis;
CC RA Zhou Z.M.;
CC RT "Cloning of a new protein kinase gene related to human testis
CC development.";
CC RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC CC -!- TISSUE SPECIFICITY: Testis specific.
CC CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC CC -!- SIMILARITY: Contains 1 Tudor domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF285599; AAK31978.1; -.
CC EMBL; AF332194; AAK3193.1; -.
CC Genew; HGNC:11407; STK31.
CC MIM; 605790; -.
CC InterPro; IPR008191; Maternal_tudor.
```


DT 01-MAY-1991 (Rel. 18, last sequence update)
 DT 10-OCT-2003 (Rel. 42, last annotation update)
 DE Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase
 DE complex (EC 2.3.1.12) (E2) (S complex, 48 kDa subunit).
 GN PDHC OR ACEC OR BSO14600.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]_TaxID=1423;
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=90368558; PubMed=1697575;
 RA Hemila H., Palva A., Paulin L., Arvidson S., Palva I.;
 RT "Secretory S complex of Bacillus subtilis: sequence analysis and
 RT identity to pyruvate dehydrogenase.";
 RL J. Bacteriol. 172:5052-5063 (1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=97124187; PubMed=8969500;
 RA Winters P., Caldwell R., Enfield L., Ferrari E.;
 RT "The ampS-nprE (124 degrees-127 degrees) region of the Bacillus
 RT subtilis 168 chromosome: sequencing of a 27 kb segment and
 RT identification of several genes in the area.";
 RL Microbiology 142:3033-3037 (1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX Caldwell R.M., Ferrari E.;
 RA "Sequence analysis of the mobA-ampS region of the Bacillus subtilis
 RT chromosome.";
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borriest L., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Gim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier P., Vassarotti A.,
 RA Viari A., Wambuit R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis.";
 RL Nature 390:249-256 (1997).
 CC -!- FUNCTION: The pyruvate dehydrogenase complex catalyzes the overall
 CC conversion of pyruvate to acetyl-CoA and CO(2). It contains
 CC multiple copies of three enzymatic components: pyruvate
 CC dehydrogenase (E1), dihydrolipoamide acetyltransferase (E2) and
 CC lipoamide dehydrogenase (E3).
 CC -!- FUNCTION: THE B.SUBTILIS PDH COMPLEX POSSESSES ALSO BRANCHED-CHAIN

CC 2-OXOACID DEHYDROGENASE (BCDH) ACTIVITY.
 CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + dihydrolipoamide = CoA + S-
 CC acetyldihydrolipoamide.
 CC -!- COFACTOR: Contains 1 covalently bound lipoyl cofactor.
 CC -!- SUBUNIT: Forms a 24-polypeptide structural core with octahedral
 CC symmetry.
 CC -!- SIMILARITY: Belongs to the 2-oxoacid dehydrogenase family.
 CC -!- SIMILARITY: Contains 1 lipoyl-binding domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M57435; AAA62683.1; --
 DR EMBL; AF012285; AAC24934.1; --
 DR EMBL; Z99111; CAB13333.1; --
 DR PIR; D36718; D36718.
 DR HSSP; P11961; 2PDE.
 DR Subtilist; BG10209; pdhC.
 DR InterPro; IPR001078; 2Oxoacid dh.
 DR InterPro; IPR000089; Biotin_lipoyl.
 DR InterPro; IPR004167; E3_binding.
 DR InterPro; IPR003016; Lipoyl_BS.
 DR Pfam; PF00198; 2-oxoacid dh; 1.
 DR Pfam; PF00364; biotin_lipoyl; 1.
 DR Pfam; PF02817; e3_binding; 1.
 DR ProDom; PD001115; 2Oxoacid dh; 1.
 DR PROSITE; PS00189; LIPOYL; 1.
 KW Glycolysis; Transferase; Acyltransferase; Lipoyl; Complete proteome.
 FT INIT MET 0 0 BY SIMILARITY.
 FT DOMAIN 1 85 LIPOYL BINDING.
 FT DOMAIN 139 181 E1/E3 BINDING.
 FT BINDING 42 42 LIPOYL.
 FT ACT_SITE 412 412 POTENTIAL.
 SQ SEQUENCE 441 AA; 47407 MW; 6B44D47879D134DF CRC64;
 Query Match 48.8%; Score 42; DB 1; Length 441;
 Best Local Similarity 46.7%; Pred. No. 17;
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 1 WKMKPRNLEEDDYL 15
 Db 21 WEVKPNDEVDV 35
 RESULT 6
 ID AGUA TRIE STANDARD; PRT; 847 AA.
 AC Q99024;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, last sequence update)
 DT 15-DEC-1998 (Rel. 37, last annotation update)
 DE Alpha-glucuronidase precursor (EC 3.2.1.139) (Alpha-glucosiduronase)
 DE (GLU1).
 OS Trichoderma reesei (Hypocrea jecorina).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Hypocreomycetidae; Hypocreales; Hypocreaeae; Hypocrea.
 OX NCBI_TaxID=51453;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=QM9414 / Rut C-30;
 RX MEDLINE=96257277; PubMed=8654984;
 RA Margolles-Clark E., Saloheimo M., Siika-Aho M., Penttilae M.;
 RT "The alpha-glucuronidase-encoding gene of Trichoderma reesei.";
 RL Gene 172:171-172 (1996).
 CC -!- FUNCTION: RELEASES 4-O-METHYLGLUCURONIC ACID FROM XYLAN.
 CC -!- CATALYTIC ACTIVITY: An alpha-D-glucuronoside + H(2)O = an alcohol
 CC + D-glucuronate.
 CC -!- SUBCELLULAR LOCATION: Secreted.

```

RT Oxodecanoyl]-L-homoserine lactone.";
RL J. Bacteriol. 179:3004-3012(1997).
CC -1- FUNCTION: REQUIRED FOR THE SYNTHESIS OF N-(3-OXODECANOYL)-L-
CC HOMOSERINE LACTONE (ODHL), AN AUTOINDUCER MOLECULE WHICH BINDS TO
CC VANR.
CC
CC -1- SIMILARITY: Belongs to the autoinducer synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U69677; AAC45212.1; -.
CC InterPro: IPR001690; Autoind synth.
CC Pfam: PF00765; Autoind synth; 1.
CC PRINTS; PR01549; AUTOINDCRSYN.
CC ProDom; PD002752; Autoind synth; 1.
CC PROSITE; PS00949; AUTOINDUCERS SYNTH; 1.
CC Quorum sensing; Autoinducer synthesis.
CC SEQUENCE 193 AA; 22127 MW; 238E760CD8701188 CRC64;
CC
CC Query Match 51.2%; Score 44; DB 1; Length 193;
CC Best Local Similarity 42.9%; Pred. No. 3.2;
CC Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
CC
CC QY 1 WKMKPRNLEEDDY 14
CC |::| | | | |
CC Db 35 WELKTRGMEYDY 48
CC
CC RESULT 4
CC PR1 SCHPO STANDARD; PRT; 539 AA.
CC ID O14293; Q9UTX5;
CC DT 16-OCT-2001 (Rel. 40, Created)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Transcription factor prr1 (Pombe response regulator 1).
CC GN PRR1 OR SPAC8C9.14.
CC OS Schizosaccharomyces pombe (Fission yeast).
CC OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
CC Schizosaccharomycetales; Schizosaccharomycetaceae;
CC Schizosaccharomycetes.
CC NCBI_TaxID=4896;
CC RN [1]
CC RP SEQUENCE FROM N.A., AND FUNCTION.
CC RX MEDLINE=99279565; PubMed=10348908;
CC RA Ohmiya R., Kato C., Yamada H., Aiba H., Mizuno T.;
CC "A fission yeast gene (prr1(+)) that encodes a response regulator
CC implicated in oxidative stress response.";
CC RL J. Biochem. 125:1061-1066(1999).
CC RN [2]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=972;
CC RX MEDLINE=21848401; PubMed=11859360;
CC RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
CC Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
CC Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
CC Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
CC Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
CC Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
CC James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
CC Mooney P., Moule S., Mungall K., Murphy I., Niblett D., Odell C.,
CC Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
CC Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
CC Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
CC Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
CC Woodward J., Volkhardt G., Aert R., Robben J., Grymonprez B.,
CC Welltjens I., Vanstreels E., Rieger M., Schaefer M., Mieller-Auer S.,
CC Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,

```

```

RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Coffeau A., Cadieu E., Dreaano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Spakovski G.V., Usery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC [3]
CC SEQUENCE OF 224-413 FROM N.A.
CC RC STRAIN=968 h90;
CC RX MEDLINE=20223868; PubMed=10759889;
CC Ding D.-Q., Tomita Y., Yamamoto A., Chikashige Y., Haraguchi T.,
CC Hiraoka Y.;
RT "Large-scale screening of intracellular protein localization in living
RT fission yeast cells by the use of a GFP-fusion genomic DNA library.";
RL Genes Cells 5:169-190(2000).
CC -1- FUNCTION: INVOLVED IN OXIDATIVE STRESS. TRANSCRIPTION FACTOR THAT
CC ACTS UPON TER1 AND CTT1.
CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
CC -1- SIMILARITY: In the N-terminal section; belongs to the HSF family.
CC -1- SIMILARITY: Contains 1 response regulatory domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB041768; BAB16722.1; -.
CC EMBL; Z99168; CAB16301.1; -.
CC EMBL; AB027943; BAA87247.1; -.
CC PIR; T39150; T39150.
CC HSP; P22121; 2HTS.
CC GeneDB Spombe; SPAC8C9.14; -.
CC InterPro; IPR000232; HSF_DNA_bind.
CC InterPro; IPR002341; HSF_ET5.
CC InterPro; IPR001789; Response reg.
CC Pfam; PF00447; HSF_DNA_bind; 1.
CC Pfam; PF00072; response_reg; 1.
CC PRINTS; PR00056; HSFDOMAIN.
CC ProDom; PD001788; HSF_DNA_bind; 1.
CC ProDom; PD000039; Response_reg; 1.
CC SMART; SM00415; HSF; 1.
CC SMART; SM00448; REC; 1.
CC PROSITE; PS00434; HSF_DOMAIN; 1.
CC PROSITE; PS0110; RESPONSE_REGULATORY; 1.
CC Transcription regulation; Sensory transduction; Nuclear protein;
CC DNA-binding; Phosphorylation.
CC KW DNA_BIND 7 111
CC FT DOMAIN 369 483
CC FT MOD_RES 418 418
CC FT PHOSPHORYLATION (BY SIMILARITY).
CC SEQUENCE 539 AA; 60045 MW; F3B4945C595B2871 CRC64;
CC
CC Query Match 50.0%; Score 43; DB 1; Length 539;
CC Best Local Similarity 60.0%; Pred. No. 14;
CC Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
CC
CC QY 1 WKMKPRNLEEDDY 15
CC |::| | | | |
CC Db 364 WKRPRIILLVEDEL 378
CC
CC RESULT 5
CC ODP2 BACSU
CC ID ODP2 BACSU STANDARD; PRT; 441 AA.
CC AC P21883;
CC DT 01-MAY-1991 (Rel. 18, Created)

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; AF186249; AAF17479.1; -;
DR EMBL; AC005053; AAC79150.1; ALT_INIT.
DR EMBL; AC004969; AAD15620.2; -;
DR EMBL; BC011802; AAH11802.1; -;
DR Genew; HGNC:11378; STEAP.
DR MIM; 604415; -;
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0005911; C:intracellular junction; TAS.
DR GO; GO:0015267; F:channel/pore class transporter activity; TAS.
KW Transmembrane; Anvigen.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 119 139 POTENTIAL.
FT TRANSMEM 164 184 POTENTIAL.
FT TRANSMEM 218 238 POTENTIAL.
FT TRANSMEM 258 278 POTENTIAL.
FT TRANSMEM 291 311 POTENTIAL.
SQ SEQUENCE 339 AA; 39851 MW; 55443AI70C870387 CRC64;

Query Match 100.0%; Score 86; DB 1; Length 339;
Best Local Similarity 100.0%; Pred. No. 4.8e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKMKPRNLEDDYL 15
DB 14 WKMKPRNLEDDYL 28
|||||

RESULT 2
RND_ECOLI STANDARD; PRT; 375 AA.
AC P09155;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ribonuclease D (EC 3.1.26.3) (RNase D).
GN RND OR B1804.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
[1]
SEQUENCE FROM N.A., AND SEQUENCE OF 1-6.
RX MEDLINE=88289400; PubMed=3041371;
RA Zhang J., Deutscher M.P.;
RT "Escherichia coli RNase D: sequencing of the rnd structural gene and
RT purification of the overexpressed protein."
RL Nucleic Acids Res. 16:6265-6278(1988).
RN [2]
SEQUENCE FROM N.A.
RP STRAIN=K12 / MG1655;
RC MEDLINE=97426617; PubMed=9278503;
RX Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Snao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [3]
SEQUENCE FROM N.A.
RP STRAIN=K12;
RC MEDLINE=97251358; PubMed=9097040;
RX Itch T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
RA Isino K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,
RA Saito N., Sempel G., Seki Y., Sivasubram S., Tagami H.,
RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horluchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome

RT corresponding to the 40.1-50.0 min region on the linkage map.";
RL DNA Res. 3:379-392(1996).
RN [4]
RP SEQUENCE OF 1-38 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=94150456; PubMed=8107670;
RA Fulda M., Heinz E., Wolter F.P.;
RT "The fadD gene of Escherichia coli K12 is located close to rnd at
RT 39.6 min of the chromosomal map and is a new member of the
RT AMP-binding protein family."
RL Mol. Gen. Genet. 242:241-249(1994).

CC -!- FUNCTION: CLEAVES MULTIMERIC TENA PRECURSOR AT THE SPACER REGION.
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X07055; CAA30098.1; -;
DR EMBL; AE000274; AAC74874.1; -;
DR EMBL; D90823; BAA15599.1; -;
DR EMBL; D90824; BAA15608.1; -;
DR EMBL; D90825; BAA15613.1; -;
DR EMBL; X70994; CAA50322.1; -;
DR PIR; S01223; NRECD.
DR ECoGene; EGI0858; rnd.
DR InterPro; IPR002562; 3_5_exonuclease.
DR InterPro; IPR002121; HRDC.
DR InterPro; IPR006292; Ribonuclease D.
DR Pfam; PF01612; 3_5_exonuclease; 1.
DR Pfam; PF00570; HRDC; 1.
DR SMART; SM00474; 3SEXOC; 1.
DR SMART; SM00341; HRDC; 1.
DR TIGRfams; TIGR01388; rnd; 1.
KW Hydrolase; Nuclease; Exonuclease; tRNA processing; Complete proteome.
SQ SEQUENCE 375 AA; 42734 MW; AC056D3683712BAE CRC64;

Query Match 52.3%; Score 45; DB 1; Length 375;
Best Local Similarity 63.6%; Pred. No. 4.4;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WKMKPRNLEE 11
DB 342 WKMKPRNLEE 352
|||||

RESULT 3
VANI VIBAN STANDARD; PRT; 193 AA.
ID VANI VIBAN
AC P74945;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Autoinducer synthesis protein vanI.
GN VANI.
OS Vibrio anguillarum (Listonella anguillarum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Listonella.
OX NCBI_TaxID=55601;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=NB10 / Serotype O1;
RC MEDLINE=97284506; PubMed=9139920;
RX Milton D.L., Hardman A., Camara M., Chhabra S.R., Bycroft B.W.,
RA Stewart G.S.A.B., Williams P.;
RT "Quorum sensing in Vibrio anguillarum: characterization of the
RT vanI/vanR locus and identification of the autoinducer N-(3-

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 13:52:02 ; Search time 4.7093 Seconds
(without alignments)
165.853 Million cell updates/sec

Title: US-10-010-667A-2_COPY_14_28
Perfect score: 86
Sequence: 1 WKVKPRNLEEDYL 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	86	100.0	339	1	STEA HUMAN
2	45	52.3	375	1	RND_ECOLI
3	44	51.2	193	1	VANI_VIBAN
4	43	50.0	539	1	PRI1_SCHPO
5	42	48.8	441	1	ODP2_BACSU
6	42	48.8	847	1	AGUA_TRIRE
7	42	48.8	1018	1	ST31_MOUSE
8	42	48.8	1019	1	ST31_HUMAN
9	40.5	47.1	639	1	HTPG_TREPA
10	40.5	47.1	1415	1	ICP4_HSVMG
11	40	46.5	36	1	RET4_CHICK
12	40	46.5	137	1	RET3_XENLA
13	40	46.5	197	1	I18B_HUMAN
14	40	46.5	549	1	SVR_ARCFU
15	40	46.5	553	1	GPT1_CANAL
16	39.5	45.9	607	1	HTPG_FUSNN
17	39	45.3	166	1	VSN1_NOCAR
18	39	45.3	347	1	AKSF_METUA
19	39	45.3	457	1	IRF7_MOUSE
20	39	45.3	639	1	SLY1_DROME
21	39	45.3	656	1	SLY1_DROME
22	39	45.3	766	1	YU20_DROME
23	39	45.3	1469	1	DP27_CABEL
24	38	44.2	142	1	OM22_HUMAN
25	38	44.2	193	1	LUXI_VIBFI
26	38	44.2	345	1	LEU3_LACLA
27	38	44.2	352	1	YG35_MYCFN
28	38	44.2	370	1	HAO1_HUMAN
29	38	44.2	427	1	ODP2_BACST
30	38	44.2	691	1	VP80_NPVAC
31	38	44.2	725	1	YGL4_YEAST
32	38	44.2	729	1	KEK1_YEAST
33	38	44.2	736	1	YBL7_YEAST

34	38	44.2	839	1	YDBB_HAEIN
35	38	44.2	925	1	RNT1_SCHPO
36	38	44.2	968	1	PKD2_HUMAN
37	38	44.2	1371	1	YOL4_CABEL
38	38	44.2	3866	1	HRX_MOUSE
39	38	44.2	3969	1	HRX_HUMAN
40	37.5	43.6	716	1	ENPL_RABIT
41	37.5	43.6	795	1	ENPL_CHICK
42	37.5	43.6	802	1	ENPL_MOUSE
43	37.5	43.6	803	1	ENPL_HUMAN
44	37.5	43.6	804	1	ENPL_BOVIN
45	37	43.0	133	1	HEMI_PYRAP

P45182	haemophilus
Q09820	schizosacch
Q13563	homo sapien
Q02331	caenorhabdi
P55200	mus musculu
Q03164	homo sapien
O18750	oryctolagus
P08110	gallus galli
P08113	mus musculu
P14625	homo sapien
Q95m18	bos taurus
P37363	pyrrhocoris

ALIGNMENTS

RESULT 1

STEAD_HUMAN STANDARD; PRT; 339 AA.

AC OSUHE8; O95034; PRT; 339 AA.

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Six transmembrane epithelial antigen of prostate.

GN STEAP OR STEAP1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20056277; PubMed=10588738;

RA Hubert R.S., Vivanco I., Chen E., Rastegar S., Leong K.,

RA Mitchell S.C., Madraswala R., Zhou Y., Kuo J., Raitano A.B.,

RA Jakobovits A., Saffran D.C., Afar D.E.H.;

RT "STEAP: a prostate-specific cell-surface antigen highly expressed in

RT human prostate tumors."

RL Proc. Natl. Acad. Sci. U.S.A. 96:14523-14528 (1999).

RN [2]

RP SEQUENCE FROM N.A.

RA Abu-Threideh J., Stoneking T., Langston Y., Maupin R.;

RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Skin;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Hellon E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Griewood J., Schmutz J., Myers R.M.,

RA Buttrick A.C., Krzywinski M.I., Skalska U., Smallos D.E.,

RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length

RT human and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -!- TISSUE SPECIFICITY: Highly expressed in prostate tumors.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

Search completed: July 12, 2004, 14:10:23
Job time : 24.1512 secs

```

RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RL through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
DR EMBL: AE015166; AAN43025.1; -.
DR GO: GO:0005737; C:cytoplasm; IEA.
DR GO: GO:0008408; F:3'-5' exonuclease activity; IEA.
DR GO: GO:0003676; F:nucleic acid binding; IEA.
DR GO: GO:0004540; F:ribonuclease activity; IEA.
DR GO: GO:0008033; P:RNA processing; IEA.
DR InterPro: IPR002562; 3.5 exonuclease.
DR InterPro: IPR02121; HRDC.
DR InterPro: IPR006292; Ribonuclease_D.
DR Pfam: PF01612; 3.5 exonuclease; 1.
DR Pfam: PF00570; HRDC; 1.
DR SMART: SM00474; 35EXOC; 1.
DR SMART: SM00341; HRDC; 1.
DR TIGRFAMs: TIGR01388; rnd; 1.
KW Complete proteome.
SQ SEQUENCE 375 AA; 42762 MW; 49FA4111C3552E34 CRC64;

Query Match 52.3%; Score 45; DB 16; Length 375;
Best Local Similarity 53.6%; Pred.No. 18;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WKMKPRNLEE 11
Db 342 WKLPQNLIPE 352

RESULT 13
Q8Q0H4 PRELIMINARY; PRT; 461 AA.
ID Q8Q0H4
AC Q8Q0H4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Conserved protein.
OS Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OC NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Goel / Go1 / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Waeber A., Baeumer S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,
RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of Methanosarcina mazei: evidence for lateral gene
RT transfer between Bacteria and Archaea."
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
DR EMBL: AE013238; AAM29858.1; -.
DR InterPro: IPR002576; Archaeal_ATPase.
DR InterPro: IPR004256; DUF234.
DR Pfam: PF01637; Archaeal_ATPase; 1.
DR Pfam: PF03008; DUF234; 1.
KW Complete proteome.
SQ SEQUENCE 461 AA; 54417 MW; 3D1AE9340AE6A9 CRC64;

Query Match 52.3%; Score 45; DB 17; Length 461;
Best Local Similarity 57.1%; Pred.No. 23;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 WKMKPRNLEEDYL 15
Db 439 KVKGRKNLEDDGFL 452

RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RL through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
DR EMBL: AE015166; AAN43025.1; -.
DR GO: GO:0005737; C:cytoplasm; IEA.
DR GO: GO:0008408; F:3'-5' exonuclease activity; IEA.
DR GO: GO:0003676; F:nucleic acid binding; IEA.
DR GO: GO:0004540; F:ribonuclease activity; IEA.
DR GO: GO:0008033; P:RNA processing; IEA.
DR InterPro: IPR002562; 3.5 exonuclease.
DR InterPro: IPR02121; HRDC.
DR InterPro: IPR006292; Ribonuclease_D.
DR Pfam: PF01612; 3.5 exonuclease; 1.
DR Pfam: PF00570; HRDC; 1.
DR SMART: SM00474; 35EXOC; 1.
DR SMART: SM00341; HRDC; 1.
DR TIGRFAMs: TIGR01388; rnd; 1.
KW Complete proteome.
SQ SEQUENCE 375 AA; 42762 MW; 49FA4111C3552E34 CRC64;

Query Match 52.3%; Score 45; DB 10; Length 2058;
Best Local Similarity 58.3%; Pred.No. 1.1e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 MKPRNLEEDDY 14
Db 496 LKKRKNISEDDY 507

RESULT 15
Q96654 PRELIMINARY; PRT; 317 AA.
ID Q96654
AC Q96654;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Activated protein kinase C receptor homolog.
OS Trypanosoma vivax.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OC NCBI_TaxID=5699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ILDat1.2;
RA Djikeng A., Majiwa P.A.;
RT "Trypanosoma vivax homologue of the receptor for activated protein
RT kinase C."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF100287; AAC72850.1; -.
DR GO: GO:0016301; F:kinase activity; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 7.
DR PRINTS: PR00320; GPROTEINRPT.
DR ProDom: PD000018; WD40; 4.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS00678; WD_REPEATS_1; 3.
DR PROSITE: PS00082; WD_REPEATS_2; 5.
DR PROSITE: PS0294; WD_REPEATS_REGION; 1.
KW Kinase; Receptor; Repeat; WD_repeat.
SQ SEQUENCE 317 AA; 34895 MW; B3CFB35D520AE31C CRC64;

Query Match 51.2%; Score 44; DB 5; Length 317;
Best Local Similarity 50.0%; Pred.No. 23;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 WKMKPRNLEEDDY 14
Db 43 WSDNPNRHAENDY 56

```

```

Query Match      52.3%; Score 45; DB 16; Length 371;
Best Local Similarity 63.6%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 WKMKPRNLEE 11
      ||:|:|:|:|
Db      338 WKLKPNLNLE 348

RESULT 10
Q8XDR7 PRELIMINARY; PRT; 375 AA.
AC Q8XDR7;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE RNase D, processes tRNA precursor.
GN RND OR Z2847 OR ECS2513.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,
RA Apodaca J., Anantharanan T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RT Nature 409:529-533 (2001).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=O157:H7 / RMD 0509952;
RC MEDLINE=21156231; PubMed=11258796;
RX Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RX Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22 (2001).
DR EMBL; AE005403; AGS6793.1; -
DR EMBL; AP002558; BAB35936.1; -
DR PIR; A99943; A99943.
DR PIR; E85791; E85791.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0008408; F:3'-5' exonuclease activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0004540; F:ribonuclease activity; IEA.
DR GO; GO:0008033; P:tRNA processing; IEA.
DR InterPro; IPR002562; 3_5_exonuclease.
DR InterPro; IPR002121; HRDC.
DR InterPro; IPR006292; Ribonuclease_D.
DR Pfam; PF01612; 3_5_exonuclease; 1.
DR Pfam; PF00570; HRDC; 1.
DR SMART; SM00341; HRDC; 1.
DR TIGRfams; TIGR01388; rnd; 1.
DR Complete proteome.
QY      1 WKMKPRNLEE 11
      ||:|:|:|:|
Db      342 WKLKPNLNLE 352

Query Match      52.3%; Score 45; DB 16; Length 375;
Best Local Similarity 63.6%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 WKMKPRNLEE 11
      ||:|:|:|:|
Db      342 WKLKPNLNLE 352

RESULT 11
Q8FGV1 PRELIMINARY; PRT; 375 AA.
AC Q8FGV1;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Ribonuclease D (EC 3.1.26.3).
GN RND OR C2208.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnerberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AE016761; AAN80667.1; -
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0008408; F:3'-5' exonuclease activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0004540; F:ribonuclease activity; IEA.
DR GO; GO:0008033; P:tRNA processing; IEA.
DR InterPro; IPR002562; 3_5_exonuclease.
DR InterPro; IPR002121; HRDC.
DR InterPro; IPR006292; Ribonuclease_D.
DR Pfam; PF01612; 3_5_exonuclease; 1.
DR Pfam; PF00570; HRDC; 1.
DR SMART; SM00341; HRDC; 1.
DR TIGRfams; TIGR01388; rnd; 1.
DR Hydrolyase; Complete proteome.
QY      1 WKMKPRNLEE 11
      ||:|:|:|:|
Db      342 WKLKPNLNLE 352

Query Match      52.3%; Score 45; DB 16; Length 375;
Best Local Similarity 63.6%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 WKMKPRNLEE 11
      ||:|:~|:~|:~|
Db      342 WKLKPNLNLE 352

RESULT 12
Q83RI1 PRELIMINARY; PRT; 375 AA.
AC Q83RI1;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE RNase D, processes tRNA precursor.
GN RND OR SFI424.
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,

```

```

Query Match          54.7%; Score 47; DB 16; Length 243;
Best Local Similarity 40.0%; Pred. No. 5.2;
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WKMKPRRNLEEDDY 15
Db 91 WKVRFQKLEDSVI 105

RESULT 6
Q8F081
ID Q8F081 PRELIMINARY; PRT; 475 AA.
AC Q8F081
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Outer membrane protein OmpA family.
GN LA3615.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; A5011517; AAN50813.1; -.
DR InterPro; IPR006664; Bac OmpA.
DR Pfam; PF00691; OmpA; 1.
DR PRINTS; PR01021; OMPADOMAIN.
DR ProDom; PD000930; OmpA/MotB; 1.
KW Complete proteome.
SQ SEQUENCE 475 AA; 55426 MW; 5394208BB7DC717B CRC64;

Query Match          54.7%; Score 47; DB 16; Length 475;
Best Local Similarity 50.0%; Pred. No. 10;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 WKMKPRRNLEEDDY 14
Db 68 WKLKFGENVELNEY 81

RESULT 7
Q83878
ID Q83878 PRELIMINARY; PRT; 208 AA.
AC Q83878
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Conserved domain protein.
GN EF0345.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacilliales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V583 / ATCC 700802;
RX MEDLINE=22550857; PubMed=12663927;
RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,
RA Daugherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,
RA Vamathevan J., Tran B., Oulton J., Hansen T., Shetty J., Khouri H.,
RA Usterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;
RT "Role of mobile DNA in the evolution of vancomycin-resistant
RT Enterococcus faecalis."
RL Science 299:2071-2074(2003).
DR EMBL; AE016947; AAC0208.1; -.
DR TIGR; EF0345; -.
KW Complete proteome.
SQ SEQUENCE 208 AA; 22905 MW; 4050BD05B0CBDA98 CRC64;

```

```

Query Match          53.5%; Score 46; DB 16; Length 208;
Best Local Similarity 61.5%; Pred. No. 6.6;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 MKKPRRNLEEDDY 15
Db 117 IKPKRYLEEGDYI 129

RESULT 8
Q9UTA5
ID Q9UTA5 PRELIMINARY; PRT; 237 AA.
AC Q9UTA5
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Conserved hypothetical protein.
GN SPAC25B8.15C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Barrell B.G., Rajandream M.A., McDougall R.C., McLean J., Harris D.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL133225; CAB61781.1; -.
DR PIR; T50202; T50202.
DR GeneDB SPombe; SPAC25B8.15C; -.
DR InterPro; IPR003827; DUF207.
DR Pfam; PF02676; DUF207; 1.
KW Hypothetical protein.
SQ SEQUENCE 237 AA; 26997 MW; 29A6CF0D0C720997 CRC64;

Query Match          52.9%; Score 45.5; DB 3; Length 237;
Best Local Similarity 66.7%; Pred. No. 9.3;
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 2 KMKPRRNLEEDDY 15
Db 220 RYKPRKLRNMDDYL 234

RESULT 9
Q7UCL0
ID Q7UCL0 PRELIMINARY; PRT; 371 AA.
AC Q7UCL0
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RNase D.
GN RND OR S1539.
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T."
RL Infect. Immun. 71:2775-2786(2003).
DR EMBL; AE016983; AAP16920.1; -.
SQ SEQUENCE 371 AA; 42225 MW; 00577552B66FEA2D CRC64;

```


01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Six-transmembrane epithelial antigen of the prostate.
 GN STEAP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; PubMed=11479226;
 RX MEDLINE=21371909; Velders M.P., Kwon E.D., Kast W.M.;
 RA Yang D., Holt G.E., Velders M.P., Kwon E.D., Kast W.M.;
 RT "Murine six-transmembrane epithelial antigen of the prostate, prostate-
 stem cell antigen, and prostate-specific membrane antigen: prostate-
 RT specific cell-surface antigens highly expressed in prostate cancer of
 RT transgenic adenocarcinoma mouse prostate mice.";
 RL Cancer Res. 61:5857-5860(2001).
 DR EMBL; AF297098; AAK83126.1; -.
 DR MGD; MGI:1917608; Steap.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR Transmembrane.
 SK SEQUENCE 339 AA; 39318 MW; 4B26A71FF559E84F CRC64;
 QW
 Query Match 73.3%; Score 63; DB 11; Length 339;
 Best Local Similarity 71.4%; Pred. No. 0.011;
 Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WKMKPRNLEEDDY 14
 Db 14 WKMKPKGNLEDDSY 27
 |||||: ||||: ||
 RESULT 3
 Q92422
 ID Q92422 PRELIMINARY; PRT; 339 AA.
 AC Q92422;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Dudulin.
 GN STEAP OR 1010001D01RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Seru V., Manivet P., Lamblin D., Vaubourdel M., Kellermann O.,
 RA Loric S.;
 RT "Prostate and non-prostate expression of dudulin, the mouse ortholog
 RT of human STEAP.";
 RL Submitted (AFR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY029584; AAK50537.1; -.
 DR MGD; MGI:1917608; Steap.
 DR SEQUENCE 339 AA; 39109 MW; 32A2C29F2E333BD0 CRC64;
 QW
 Query Match 73.3%; Score 63; DB 11; Length 339;
 Best Local Similarity 71.4%; Pred. No. 0.011;
 Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WKMKPRNLEEDDY 14
 Db 14 WKMKPKGNLEDDSY 27
 |||||: ||||: ||
 RESULT 4
 Q9CWR7
 ID Q9CWR7 PRELIMINARY; PRT; 339 AA.
 AC Q9CWR7;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE 2410007B19RIK protein.

GN STEAP OR 2410007B19RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryonic stem cells;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 DR EMBL; AK010437; BAB26938.1; -.
 DR MGD; MGI:1917608; Steap.
 DR SEQUENCE 339 AA; 39264 MW; 3F7AB9C7520F0968 CRC64;
 QY 1 WKMKPRNLEEDDY 14
 Db 14 WKMKPKGNLEDDSY 27
 |||||: ||||: ||
 RESULT 5
 Q87NG5
 ID Q87NG5 PRELIMINARY; PRT; 243 AA.
 AC Q87NG5;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative DNA polymerase III, epsilon subunit.
 GN VPI903.
 OS Vibrio parahaemolyticus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=670;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RIMD 2210633 / Serotype O3:K6;
 RX MEDLINE=22508454; PubMed=12620739;
 RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
 RA Iijima Y., Neijima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
 RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
 RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
 RT distinct from that of V. cholerae.";
 RL Lancet 361:743-749(2003).
 DR EMBL; AP005079; BAC60166.1; -.
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0004527; F:exonuclease activity; IEA.
 DR InterPro; IPR006055; Exonuclease.
 DR Pfam; PF00929; Exonuclease; 1.
 KW Complete proteome.
 SQ SEQUENCE 243 AA; 27903 MW; 6696A2E2A0291751 CRC64;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:00:58 ; Search time 22.1512 Seconds
(without alignments)
213.658 Million cell updates/sec

Title: US-10-010-667a-2_COPY_14_28

Perfect score: 86

Sequence: 1 WKMKPRNLEDDYL 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67.5	78.5	338	6	Q9GL50 sus scrofa
2	63	73.3	339	11	Q924J9
3	63	73.3	339	11	Q924Z2
4	63	73.3	339	11	Q9CWR7
5	47	54.7	243	16	Q87NG5
6	47	54.7	475	16	Q8F081
7	46	53.5	208	16	Q838T8
8	45.5	52.9	237	3	Q9UTA5
9	45	52.3	371	16	Q7UCLO
10	45	52.3	375	16	Q8XDR7
11	45	52.3	375	16	Q8FGV1
12	45	52.3	375	16	Q83R11
13	45	52.3	461	17	Q8Q0H4
14	45	52.3	2058	10	Q9LGG0
15	44	51.2	317	5	Q96654
16	44	51.2	532	10	O80941

17	44	51.2	659	5	O76517
18	44	51.2	773	2	Q91UQ5
19	43	50.0	53	16	Q8EVE2
20	43	50.0	172	16	Q81MG8
21	43	50.0	312	5	Q17557
22	43	50.0	348	8	Q8M3F6
23	43	50.0	348	8	Q8M3F2
24	43	50.0	348	8	Q8M3F1
25	43	50.0	398	10	Q94HS8
26	43	50.0	398	10	Q7XGV6
27	43	50.0	491	10	Q7XSE9
28	43	50.0	679	13	Q98856
29	43	50.0	2040	12	Q8B0U2
30	42	48.8	121	10	Q7XXZ3
31	42	48.8	124	16	Q928F4
32	42	48.8	146	5	Q86S8
33	42	48.8	316	5	O61129
34	42	48.8	331	13	Q7ZW96
35	42	48.8	384	10	Q9LXJ7
36	42	48.8	416	10	Q9CGH1
37	42	48.8	441	10	Q9FJ11
38	42	48.8	451	10	Q9LXJ6
39	42	48.8	564	16	O81G72
40	42	48.8	1065	2	Q9AHK8
41	42	48.8	1417	5	Q9VSE1
42	42	48.8	1858	16	Q8UIZ6
43	41.5	48.3	435	5	Q9TYG6
44	41.5	48.3	435	5	Q9TZR3
45	41.5	48.3	435	5	Q9GQW8

ALIGNMENTS

RESULT 1

Q9GL50 PRELIMINARY; PRT; 338 AA.

AC Q9GL50;

DT 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE Six transmembrane endothelial antigen of PAEC.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A.

RA Nagasaka T., Boulday G., Coupel S., Coulon F., Tesson L., Heslan J.-M., Soullion J.-P., Charreau B.; "Differential gene expression in endothelial cells during TNF-alpha- and LPS-mediated activation.";

RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF319659; AAG33868.1; -.

DR GO; GO:0016021; C:integral to membrane; IEA.

KW Transmembrane.

SQ SEQUENCE 338 AA; 39918 MW; ED490E86E067A32B CRC64;

Query Match 78.5%; Score 67.5; DB 6; Length 338;

Best Local Similarity 86.7%; Pred. No. 0.0018;

Matches 13; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 WKMKPRNLEDDYL 15

Db 14 WKMKPRNLEDDYL 27

RESULT 2

Q924J9 PRELIMINARY; PRT; 339 AA.

ID Q924J9

AC Q924J9;

DT 01-DEC-2001 (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

ABU98390
 ID ABU98390 standard; protein; 258 AA.
 XX
 AC ABU98390;
 XX
 DT 31-JUL-2003 (first entry)
 XX
 DE Novel human gene STEAP-1 variant 8.
 XX
 KW STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
 KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
 KW vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO2003022995-A2.
 XX
 PD 20-MAR-2003.
 XX
 PF 06-SEP-2002; 2002WO-US028371.
 XX
 PR 06-SEP-2001; 2001US-0317840P.
 PR 05-APR-2002; 2002US-0370387P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 PI Paris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
 XX
 DR WPI; 2003-313240/30.
 DR N-PSDB; ACD02604.
 XX
 PT New composition comprising a substance that modulates the status of a
 PT STEAP-1-related protein, useful for treating and detecting cancer.
 XX
 PS Example 2; Fig 2H; 248pp; English.
 XX
 CC The invention describes a composition comprising a substance that
 CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
 CC any of the 15 sequences of 259 amino acids, given in the specification,
 CC or a molecule that is modulated by the protein, where the status of the
 CC cell that expresses the protein is modulated. The compositions, proteins,
 CC polynucleotides and methods are useful for treating and detecting cancer.
 CC The STEAP-1-related proteins are useful for generating cancer vaccines.
 CC The polynucleotides are useful as tools for delineating, with greater
 CC precision, cytogenetic abnormalities in the chromosomal region that
 CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
 CC the amino acid sequence of a variant of human six transmembrane
 CC epithelial antigen of the prostate or STEAP-1
 XX
 SQ Sequence 258 AA;
 Query Match 100.0%; Score 192; DB 6; Length 258;
 Best Local Similarity 100.0%; Pred. No. 3.9e-18;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RRSYRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34
 DB 185 RRSYRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 218

Search completed: July 12, 2004, 14:07:23
 Job time : 75.7209 secs

CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
CC any of the 15 sequences of 259 amino acids, given in the specification,
CC or a molecule that is modulated by the protein, where the status of the
CC cell that expresses the protein is modulated. The compositions, proteins,
CC polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
XX
SQ Sequence 258 AA;

Query Match 100.0%; Score 192; DB 6; Length 258;
Best Local Similarity 100.0%; Pred. No. 3.9e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 34
185 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 218

RESULT 13
ABU98387

ID ABU98387 standard; protein; 258 AA.

XX AC ABU98387;

DT 31-JUL-2003 (first entry)

DE Novel human gene STEAP-1 variant 5.

XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
XX cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX vaccine.

OS Homo sapiens.

XX WO2003022995-A2.

XX 20-MAR-2003.

PF 06-SEP-2002; 2002WO-US028371.

XX 06-SEP-2001; 2001US-0317840P.

PR 05-APR-2002; 2002US-0370387P.

XX (AGEN-) AGENSYS INC.

XX Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;

XX WPI; 2003-313240/30.

XX N-PSDB; ACDC02601.

PT New composition comprising a substance that modulates the status of a
XX STEAP-1-related protein, useful for treating and detecting cancer.

XX Example 2; Fig 2E; 248pp; English.

CC The invention describes a composition comprising a substance that
CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
CC any of the 15 sequences of 259 amino acids, given in the specification,
CC or a molecule that is modulated by the protein, where the status of the
CC cell that expresses the protein is modulated. The compositions, proteins,
CC polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
XX

SQ Sequence 258 AA;

Query Match 100.0%; Score 192; DB 6; Length 258;
Best Local Similarity 100.0%; Pred. No. 3.9e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 34
185 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 218

RESULT 14
ABU98393

ID ABU98393 standard; protein; 258 AA.

XX AC ABU98393;

DT 31-JUL-2003 (first entry)

DE Novel human gene STEAP-1 variant 11.

XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
XX cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX vaccine.

OS Homo sapiens.

XX WO2003022995-A2.

XX 20-MAR-2003.

PF 06-SEP-2002; 2002WO-US028371.

XX 06-SEP-2001; 2001US-0317840P.

PR 05-APR-2002; 2002US-0370387P.

XX (AGEN-) AGENSYS INC.

XX Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;

XX WPI; 2003-313240/30.

XX N-PSDB; ACDC02607.

PT New composition comprising a substance that modulates the status of a
XX STEAP-1-related protein, useful for treating and detecting cancer.

XX Example 2; Fig 2K; 248pp; English.

CC The invention describes a composition comprising a substance that
CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
CC any of the 15 sequences of 259 amino acids, given in the specification,
CC or a molecule that is modulated by the protein, where the status of the
CC cell that expresses the protein is modulated. The compositions, proteins,
CC polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
XX

SQ Sequence 258 AA;

Query Match 100.0%; Score 192; DB 6; Length 258;
Best Local Similarity 100.0%; Pred. No. 3.9e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 34
185 RRSRYKLLNWAYQQVQONKEDAWIEHDVVRMEI 218

RESULT 15

PA (AGEN-) AGENSYS INC.
 XX Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
 PI WPI; 2003-313240/30.
 XX
 DR New composition comprising a substance that modulates the status of a
 PT STEAP-1-related protein, useful for treating and detecting cancer.
 XX
 PS Example 53; Page 169; 248pp; English.
 XX
 CC The invention describes a composition comprising a substance that
 CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
 CC any of the 15 sequences of 259 amino acids, given in the specification,
 CC or a molecule that is modulated by the protein, where the status of the
 CC cell that expresses the protein is modulated. The compositions, proteins,
 CC polynucleotides and methods are useful for treating and detecting cancer.
 CC The STEAP-1-related proteins are useful as tools for delineating, with greater
 CC precision, cytogenetic abnormalities in the chromosomal region that
 CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
 CC the amino acid sequence of a variant of human six transmembrane
 CC epithelial antigen of the prostate or STEAP-1
 XX
 PS Sequence 258 AA;
 CC
 CC Query Match 100.0%; Score 192; DB 6; Length 258;
 CC Best Local Similarity 100.0%; Pred. No. 3.9e-18; Mismatches 0; Gaps 0;
 CC Matches 34; Conservative 0; Indels 0; Gaps 0;
 QY 1 RRSRYKLLNWAYQVQVQNKEDAWIEHDVWRMEI 34
 DB 185 RRSRYKLLNWAYQVQVQNKEDAWIEHDVWRMEI 218
 RESULT 12
 ABU98392
 ID ABU98392 standard; protein; 258 AA.
 AC ABU98392;
 XX
 DT 31-JUL-2003 (first entry)
 XX
 DE Novel human gene STEAP-1 variant 10.
 XX
 KW STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
 KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
 KW vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO2003022995-A2.
 XX
 PD 20-MAR-2003.
 XX
 PF 06-SEP-2002; 2002WO-US028371.
 XX
 PR 06-SEP-2001; 2001US-0317840P.
 PR 05-APR-2002; 2002US-0370387P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 XX Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
 PI WPI; 2003-313240/30.
 DR N-PSDB; ACD02606.
 XX
 CC New composition comprising a substance that modulates the status of a
 PT STEAP-1-related protein, useful for treating and detecting cancer.
 XX
 PS Example 2; Fig 2G; 248pp; English.
 XX
 CC The invention describes a composition comprising a substance that
 CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
 CC any of the 15 sequences of 259 amino acids, given in the specification,
 CC or a molecule that is modulated by the protein, where the status of the
 CC cell that expresses the protein is modulated. The compositions, proteins,
 CC polynucleotides and methods are useful for treating and detecting cancer.
 CC The STEAP-1-related proteins are useful as tools for delineating, with greater
 CC precision, cytogenetic abnormalities in the chromosomal region that
 CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
 CC the amino acid sequence of a variant of human six transmembrane
 CC epithelial antigen of the prostate or STEAP-1
 XX
 PS Sequence 258 AA;
 CC
 CC Query Match 100.0%; Score 192; DB 6; Length 258;
 CC Best Local Similarity 100.0%; Pred. No. 3.9e-18; Mismatches 0; Gaps 0;
 CC Matches 34; Conservative 0; Indels 0; Gaps 0;
 QY 1 RRSRYKLLNWAYQVQVQNKEDAWIEHDVWRMEI 34
 DB 185 RRSRYKLLNWAYQVQVQNKEDAWIEHDVWRMEI 218
 RESULT 11
 ABU98424
 ID ABU98424 standard; protein; 258 AA.
 AC ABU98424;
 XX
 DT 31-JUL-2003 (first entry)
 XX
 DE STEAP-1 variant 8FLD4 v.2 #1.
 XX
 KW STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
 KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
 KW vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO2003022995-A2.
 XX
 PD 20-MAR-2003.
 XX
 PF 06-SEP-2002; 2002WO-US028371.
 XX
 PR 06-SEP-2001; 2001US-0317840P.
 PR 05-APR-2002; 2002US-0370387P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 XX Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
 PI WPI; 2003-313240/30.
 DR N-PSDB; ACD02606.
 XX
 CC New composition comprising a substance that modulates the status of a
 PT STEAP-1-related protein, useful for treating and detecting cancer.
 XX
 PS Example 2; Fig 2G; 248pp; English.
 XX
 CC The invention describes a composition comprising a substance that
 CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
 CC any of the 15 sequences of 259 amino acids, given in the specification,
 CC or a molecule that is modulated by the protein, where the status of the
 CC cell that expresses the protein is modulated. The compositions, proteins,
 CC polynucleotides and methods are useful for treating and detecting cancer.
 CC The STEAP-1-related proteins are useful as tools for delineating, with greater
 CC precision, cytogenetic abnormalities in the chromosomal region that
 CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
 CC the amino acid sequence of a variant of human six transmembrane
 CC epithelial antigen of the prostate or STEAP-1
 XX
 PS Sequence 258 AA;
 CC
 CC Query Match 100.0%; Score 192; DB 6; Length 258;
 CC Best Local Similarity 100.0%; Pred. No. 3.9e-18; Mismatches 0; Gaps 0;
 CC Matches 34; Conservative 0; Indels 0; Gaps 0;
 QY 1 RRSRYKLLNWAYQVQVQNKEDAWIEHDVWRMEI 34
 DB 185 RRSRYKLLNWAYQVQVQNKEDAWIEHDVWRMEI 218

XX Homo sapiens.
 OS WO2003022995-A2.
 XX
 PN 20-MAR-2003.
 XX
 PD 06-SEP-2002; 2002WO-US028371.
 XX
 PF 06-SEP-2001; 2001US-0317840P.
 PR 05-APR-2002; 2002US-0370387P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 XX Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
 PI WPI; 2003-313240/30.
 DR N-PSDB; ACD02606.
 XX
 CC New composition comprising a substance that modulates the status of a
 PT STEAP-1-related protein, useful for treating and detecting cancer.
 XX
 PS Example 2; Fig 2G; 248pp; English.
 XX
 CC The invention describes a composition comprising a substance that
 CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
 CC any of the 15 sequences of 259 amino acids, given in the specification,
 CC or a molecule that is modulated by the protein, where the status of the
 CC cell that expresses the protein is modulated. The compositions, proteins,
 CC polynucleotides and methods are useful for treating and detecting cancer.
 CC The STEAP-1-related proteins are useful as tools for delineating, with greater
 CC precision, cytogenetic abnormalities in the chromosomal region that
 CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
 CC the amino acid sequence of a variant of human six transmembrane
 CC epithelial antigen of the prostate or STEAP-1
 XX
 PS Sequence 258 AA;
 CC
 CC Query Match 100.0%; Score 192; DB 6; Length 258;
 CC Best Local Similarity 100.0%; Pred. No. 3.9e-18; Mismatches 0; Gaps 0;
 CC Matches 34; Conservative 0; Indels 0; Gaps 0;
 QY 1 RRSRYKLLNWAYQVQVQNKEDAWIEHDVWRMEI 34
 DB 185 RRSRYKLLNWAYQVQVQNKEDAWIEHDVWRMEI 218
 RESULT 11
 ABU98424
 ID ABU98424 standard; protein; 258 AA.
 AC ABU98424;
 XX
 DT 31-JUL-2003 (first entry)
 XX
 DE STEAP-1 variant 8FLD4 v.2 #1.
 XX
 KW STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
 KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
 KW vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO2003022995-A2.
 XX
 PD 20-MAR-2003.
 XX
 PF 06-SEP-2002; 2002WO-US028371.
 XX
 PR 06-SEP-2001; 2001US-0317840P.
 PR 05-APR-2002; 2002US-0370387P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 XX Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
 PI WPI; 2003-313240/30.
 DR N-PSDB; ACD02606.
 XX
 CC New composition comprising a substance that modulates the status of a
 PT STEAP-1-related protein, useful for treating and detecting cancer.
 XX
 PS Example 2; Fig 2G; 248pp; English.
 XX
 CC The invention describes a composition comprising a substance that
 CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
 CC any of the 15 sequences of 259 amino acids, given in the specification,
 CC or a molecule that is modulated by the protein, where the status of the
 CC cell that expresses the protein is modulated. The compositions, proteins,
 CC polynucleotides and methods are useful for treating and detecting cancer.
 CC The STEAP-1-related proteins are useful as tools for delineating, with greater
 CC precision, cytogenetic abnormalities in the chromosomal region that
 CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
 CC the amino acid sequence of a variant of human six transmembrane
 CC epithelial antigen of the prostate or STEAP-1
 XX
 PS Sequence 258 AA;
 CC
 CC Query Match 100.0%; Score 192; DB 6; Length 258;
 CC Best Local Similarity 100.0%; Pred. No. 3.9e-18; Mismatches 0; Gaps 0;
 CC Matches 34; Conservative 0; Indels 0; Gaps 0;
 QY 1 RRSRYKLLNWAYQVQVQNKEDAWIEHDVWRMEI 34
 DB 185 RRSRYKLLNWAYQVQVQNKEDAWIEHDVWRMEI 218

XX SQ Sequence 255 AA;

Query Match 100.0%; Score 192; DB 6; Length 255;
Best Local Similarity 100.0%; Pred. No. 3.9e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34
185 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 218

Db

RESULT 8

ABU98428

ID ABU98428 standard; protein; 255 AA.

XX AC ABU98428;

XX AC

DT 31-JUL-2003 (first entry)

XX 31-JUL-2003 (first entry)

DE STEAP-1 variant 8P1D4 v.2 #3.

XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
vaccine.

KW

KW

OS Homo sapiens.

XX

PN WO2003022995-A2.

XX WO2003022995-A2.

PD 20-MAR-2003.

XX 20-MAR-2003.

PF 06-SEP-2002; 2002WO-US028371.

XX 06-SEP-2002; 2002WO-US028371.

PR 06-SEP-2001; 2001US-0317840P.

XX 06-SEP-2001; 2001US-0317840P.

PR 05-APR-2002; 2002US-0370387P.

XX 05-APR-2002; 2002US-0370387P.

XX (AGEN-) AGENSYS INC.

PA (AGEN-) AGENSYS INC.

XX

PI Faris M, Ge W, Raitano AB, Challita-Bid PM, Jakobovits A;

XX WPI; 2003-313240/30.

DR N-PSDB; AC002609.

XX

PT New composition comprising a substance that modulates the status of a
STEAP-1-related protein, useful for treating and detecting cancer.

XX Example 53; Page 172; 248pp; English.

CC The invention describes a composition comprising a substance that
modulates the status of a protein (I) of 340 or 283 amino acids, or of
any of the 15 sequences of 259 amino acids, given in the specification,
or a molecule that is modulated by the protein, where the status of the
cell that expresses the protein is modulated. The compositions, proteins,
polynucleotides and methods are useful for treating and detecting cancer.
The STEAP-1-related proteins are useful for generating cancer vaccines.
The polynucleotides are useful as tools for delineating, with greater
precision, cytogenetic abnormalities in the chromosomal region that
encodes STEAP-1 that may contribute to the malignant phenotype. This is
the amino acid sequence of a variant of human six transmembrane
epithelial antigen of the prostate or STEAP-1

XX SQ Sequence 255 AA;

Query Match 100.0%; Score 192; DB 6; Length 255;
Best Local Similarity 100.0%; Pred. No. 3.9e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 34
185 RRSRYKLLNWAYQQVQONKEDAWIEHDVWRMEI 218

Db

RESULT 9

ABU98389

ID ABU98389 standard; protein; 258 AA.

XX AC ABU98389;

XX AC

DT 31-JUL-2003 (first entry)

XX 31-JUL-2003 (first entry)

DE Novel human gene STEAP-1 variant 7.

XX Novel human gene STEAP-1 variant 7.

KW STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
vaccine.

KW

PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 197
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)...(28)
US-09-485-632B-6

Query Match 46.5%; Score 40; DB 4; Length 197;
Best Local Similarity 46.7%; Pred. No. 36;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 WKMKPRNLEEDDY 15
DB 167 WYSPRRLQOEEL 181

RESULT 7
US-09-247-373B-54
Sequence 54, Application US/09247373B
Patent No. 6168954
GENERAL INFORMATION:
APPLICANT: MCGONIGLE, BRIAN
APPLICANT: O'KEEFE, DANIEL
TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES
FILE REFERENCE: CL-1108-A
CURRENT APPLICATION NUMBER: US/09/247,373B
CURRENT FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 08/924,747
PRIOR FILING DATE: 1997-09-05
NUMBER OF SEQ ID NOS: 56
SOFTWARE: Microsoft Office 97
SEQ ID NO 54
LENGTH: 221
TYPE: PRT
ORGANISM: SOYBEAN
US-09-247-373B-54

Query Match 46.5%; Score 40; DB 3; Length 221;
Best Local Similarity 44.4%; Pred. No. 41;
Matches 8; Conservative 3; Mismatches 3; Indels 4; Gaps 1;

QY 1 WKMK---PRNLEEDDY 14
DB 18 WYWKLDIPYNTBEDRY 35

RESULT 8
US-09-252-991A-31334
Sequence 31334, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196, 136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 31334
LENGTH: 573
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31334

Query Match 46.5%; Score 40; DB 4; Length 573;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 PRNLEEDDY 14
DB 427 PRPNDEDDY 436

RESULT 9
US-09-621-976-5168
Sequence 5168, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 5168
LENGTH: 103
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: UNSURE
LOCATION: 68
OTHER INFORMATION: Xaa = Asp,Tyr
US-09-621-976-5168

Query Match 44.2%; Score 38; DB 4; Length 103;
Best Local Similarity 70.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 KPRNLEEDD 13
DB 27 KPEELEEDD 36

RESULT 10
US-09-582-379-4
Sequence 4, Application US/09582379
Patent No. 6492176
GENERAL INFORMATION:
APPLICANT: K.U. Leuven Research & Development
TITLE OF INVENTION: Increase of No. 6492176ule number and nitrogen fixation in
FILE REFERENCE: 001252
CURRENT APPLICATION NUMBER: US/09/582,379
CURRENT FILING DATE: 2000-10-10
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 193
TYPE: PRT
ORGANISM: Vibri0 fischeri
US-09-582-379-4

Query Match 44.2%; Score 38; DB 4; Length 193;
Best Local Similarity 42.9%; Pred. No. 77;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 WKMKPRNLEEDDY 14
DB 35 WDLVVENNLSEDEX 48

RESULT 11
US-08-403-866-2
Sequence 2, Application US/08403866

Patent No. 5643779
GENERAL INFORMATION:
APPLICANT: Ehrlich, Stanislaw
APPLICANT: Godon, Jean-Jacques
APPLICANT: Renault, Pierre
TITLE OF INVENTION: Nucleic acid coding for an alpha-acetolactate
TITLE OF INVENTION: synthase from Lactococcus and its applications
NUMBER OF SEQUENCES: 16
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/403,866
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30, 727
REFERENCE/DOCKET NUMBER: 20747/30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1600
TELEX: 978450 (WDT)
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 345 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Lactococcus lactis subsp. lactis
INDIVIDUAL ISOLATE: LEUB
US-08-403-866-2

Query Match 44.2%; Score 38; DB 1; Length 345;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 KMKPRRLIEDDTL 15
DB 114 KLSPIRVNVDPL 127

RESULT 12
US-09-252-991A-29449
Sequence 29449, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 29449
LENGTH: 464
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
FEATURE:
NAME/KEY: UNSURE
LOCATION: (227), (276)
OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-29449

Query Match 44.2%; Score 38; DB 4; Length 464;

Best Local Similarity 60.0%; Pred. No. 2e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 PRRLIEDDTL 14
DB 90 PRRLDRDDH 99

RESULT 13
US-09-134-000C-4765
Sequence 4765, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: Patentin version 3.1
SEQ ID NO 4765
LENGTH: 540
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-134-000C-4765

Query Match 44.2%; Score 38; DB 4; Length 540;
Best Local Similarity 46.7%; Pred. No. 2.4e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 WKKPRRLIEDDTL 15
DB 23 WFKVPGDTINEDDTL 37

RESULT 14
US-09-107-532A-6422
Sequence 6422, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 893-5007

TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 6422:
SEQUENCE CHARACTERISTICS:
LENGTH: 555 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1..555
SEQUENCE DESCRIPTION: SEQ ID NO: 6422:
US-09-107-532A-6422

Query Match 44.2%; Score 38; DB 4; Length 555;
Best Local Similarity 46.7%; Pred. No. 2.4e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 WKMKPRRLLEDDYL 15
| : | | : | | | |
Db 30 WPKRGDTINEDDTL 44

RESULT 15
US-08-651-999A-7
Sequence 7, Application us/08651999A
Patent No. 6031088
GENERAL INFORMATION:
APPLICANT: Stefan Somlo and Toshio Mochizuki
TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE PKD2 GENE
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: AMSTER, ROTHSTEIN & EBENSTEIN
STREET: 90 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH 1.44 MB STORAGE
MEDIUM TYPE: DISKETTE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/651,999A
FILING DATE: MAY 23, 1996
ATTORNEY/AGENT INFORMATION:
NAME: ELIZABETH A. BOGOSIAN
REGISTRATION NUMBER: 39,911
REFERENCE/DOCKET NUMBER: 96700/395
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 286-0854
TELEFAX: (212) 286-0854 or 286-0082
TELEX: TWX 710-581-4766
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 968 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: peptide
HYPOTHETICAL: NO
US-08-651-999A-7

Query March 44.2%; Score 38; DB 3; Length 968;
Best Local Similarity 46.2%; Pred. No. 4.5e+02;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 WKMKPRRLLEDD 13
| : | | : | | | |
Db 293 WKMPNOTADN 305

Search completed: July 12, 2004, 14:12:23
Job time : 10.2442 secs

This page blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: July 12, 2004, 13:51:07 ; Search time 32.9651 Seconds
(without alignments)
128.567 Million cell updates/sec

Title: US-10-010-667a-2_COPY_14_28
Perfect score: 86
Sequence: 1 WKMKPRRLIEDDYLL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 segs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	86	100.0	15	3	AAY58201	AAY58201 Human STR
2	86	100.0	15	4	AAE02784	AAE02784 Immunogen
3	86	100.0	66	2	AAY12305	AAY12305 Human 5'
4	86	100.0	109	2	AAY12304	AAY12304 Human 5'
5	86	100.0	254	6	ABU98426	ABU98426 STEAP-1 v
6	86	100.0	254	6	ABU98425	ABU98425 STEAP-1 v
7	86	100.0	255	6	ABU98429	ABU98429 STEAP-1 v
8	86	100.0	255	6	ABU98428	ABU98428 STEAP-1 v
9	86	100.0	258	6	ABU98395	ABU98395 Novel hum
10	86	100.0	258	6	ABU98389	ABU98389 Novel hum
11	86	100.0	258	6	ABU98424	ABU98424 STEAP-1 v
12	86	100.0	258	6	ABU98392	ABU98392 Novel hum
13	86	100.0	258	6	ABU98387	ABU98387 Novel hum
14	86	100.0	258	6	ABU98393	ABU98393 Novel hum
15	86	100.0	258	6	ABU98390	ABU98390 Novel hum
16	86	100.0	258	6	ABU98391	ABU98391 Novel hum
17	86	100.0	258	6	ABU98384	ABU98384 Novel hum
18	86	100.0	258	6	ABU98431	ABU98431 STEAP-1 v
19	86	100.0	258	6	ABU98396	ABU98396 Novel hum
20	86	100.0	258	6	ABU98397	ABU98397 Novel hum
21	86	100.0	258	6	ABU98394	ABU98394 Novel hum
22	86	100.0	258	6	ABU98388	ABU98388 Novel hum
23	86	100.0	258	6	ABU98386	ABU98386 Novel hum
24	86	100.0	258	6	ABU98399	ABU98399 Novel hum
25	86	100.0	258	6	ABU98398	ABU98398 Novel hum

ALIGNMENTS

26	86	100.0	282	6	ABU98432	ABU98432 STEAP-1 v
27	86	100.0	282	6	ABU98427	ABU98427 STEAP-1 v
28	86	100.0	282	6	ABU98385	ABU98385 Novel hum
29	86	100.0	339	3	AAY58194	AAY58194 Human STR
30	86	100.0	339	3	AAW01282	AAW01282 P789P anti
31	86	100.0	339	4	AAU69927	AAU69927 Human pro
32	86	100.0	339	4	AAW78845	AAW78845 Human pro
33	86	100.0	339	4	ABU71818	ABU71818 Prostate
34	86	100.0	339	5	ABU95387	ABU95387 Human P78
35	86	100.0	339	5	ABG61813	ABG61813 Prostate
36	86	100.0	339	6	ABU98383	ABU98383 Novel hum
37	86	100.0	339	6	ABU798414	ABU798414 STEAP-1 v
38	86	100.0	339	6	ABU98430	ABU98430 STEAP-1 v
39	86	100.0	339	6	ABR54499	ABR54499 Prostate
40	86	100.0	339	7	ABU63313	ABU63313 Human six
41	86	100.0	339	7	ADB73573	ADB73573 Prostate
42	86	100.0	339	7	ADB14329	ADB14329 Human pro
43	86	100.0	375	4	AAW79829	AAW79829 Human pro
44	86	100.0	375	4	AAE02780	AAE02780 Human six
45	70	81.4	95	2	AAY11840	AAY11840 Human 5'

RESULT 1
ID AAY58201 standard; peptide; 15 AA.
AC AAY58201;
XX
DT 14-MAR-2000 (first entry)
DE Human STRAP-1 peptide 1, corresponding to STRAP-1 residues 14-28.

XX Serpentine transmembrane antigen of the prostate; STRAP-1; prostate;
XX transmembrane domain; type IIa membrane protein; expression; cancer;
KW prostate cancer; bladder cancer; colon cancer; pancreatic cancer;
KW ovarian cancer; tumour antigen; immunisation; immune response; cellular;
KW humoral; anticancer vaccine; antibody; detection; diagnosis; prognosis;
KW monitoring; susceptibility; therapeutic inhibitor; drug targeting;
KW recombinant protein.
XX

OS Synthetic.
OS Homo sapiens.
XX

PN WO962941-A2.
PD 09-DEC-1999.

XX 01-JUN-1999; 99WO-US012157.
PF

XX 01-JUN-1998; 98US-0087520P.
PR 30-JUN-1998; 98US-0091183P.

XX (UROC-) UROGENESYS INC.
PA (AFAR/) AFAR D E.

PA (HUBE/) HUBERT R S.
PA (LEON/) LEON K.

PA (RAIT/) RAITANO A B.
PA (SAFF/) SAFFRAN D C.

XX Afer DE, Hubert RS, Leong K, Raitano AB, Saffran DC;
FI WPI; 2000-072832/06.

XX Novel proteins useful as diagnostic markers and therapeutic targets,
PT particularly for prostatic cancer.
XX

XX Example 3C; Page 41; 83pp; English.

XX This sequence represents a synthetic peptide (peptide 1) corresponding to
CC residues 14-28 of STRAP-1 (serpentine transmembrane antigen of the

prostate, AAY8194). This peptide was used to raise polyclonal anti-STRAP-1 N-terminus antibodies in sheep. The antibodies were used to probe Western blots of cell lysates from prostate and non-prostate cancer cell lines, and in immunohistochemical analysis of STRAP-1 expression. STRAP-1 is the prototype member of the STRAP Family of proteins (AAY58194-Y58197) which exhibit a high degree of structural conservation, but which show no significant structural homology to known human proteins. The STRAP-1 gene has been localised to chromosome 7p22. STRAP-1 is thought to be a type IIIa membrane protein and is expressed predominantly in prostate cells in normal human tissues. Structurally, STRAP-1 is a 339 amino acid protein characterised by six transmembrane domains and intracellular N- and C-termini, suggesting that it folds in a "serpentine" manner into three extracellular and two intracellular loops. STRAP-1 mRNA and protein expression is maintained at high levels and throughout all stages of prostate cancer. STRAP-1 mRNA and/or protein is also overexpressed in certain other cancers, including bladder, colon, pancreatic and ovarian cancer. The function of the STRAP proteins is not known. They may be ion channels (from the presence of six transmembrane domains, a feature which is shared by certain ion channels) or gap-junction proteins (from immunohistochemical staining). STRAP-1 and STRAP-2 are cell-surface tumour antigens. Immunisation with a STRAP protein induces cellular and humoral immune responses against STRAP-expressing cells. STRAP proteins may be used to identify specific-binding agents, to produce anticancer vaccines and to generate specific antibodies. The antibodies may be used for detection, prognosis, and monitoring of cancers (or susceptibility to cancer), as therapeutic inhibitors or to target therapeutic agents to their site of action. STRAP nucleic acids may be used for recombinant protein production, as diagnostic and prognostic reagents, for identifying STRAP-expressing cells for screening inhibitors of STRAP expression and for therapeutic modulation/inhibition of STRAP expression. Since high levels of STRAP proteins are exposed on the cell surface, they are easily targeted by systemically administered agents, and because they are expressed mainly on prostatic epithelial cells, agents targeted to them should have minimal side effects on other tissues

Sequence 15 AA:

Query Match 100.0%; Score 86; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 8.3e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKMKPRRNLEDDYL 15
Db 1 WKMKPRRNLEDDYL 15

RESULT 2
AAE02784

ID AAE02784 standard; peptide; 15 AA.

AC AAE02784;

DT 06-AUG-2001 (first entry)

DE Immunogenic peptide #1 of human STRAP-1.

KW Human; cytostatic; antiproliferative; vaccine; gene therapy;
KW six transmembrane epithelial antigen of the prostate-1; STEAP-1;
KW chromosome 7p22.3; cancer; prostate; colon; bladder; pancreatic; ovarian;
KW lung; serpentine transmembrane antigen; immunogen.

OS Homo sapiens.

PN WO200140276-A2.

PD 07-JUN-2001.

PF 06-DEC-2000; 2000WO-US033040.

PR 06-DEC-1999; 99US-00455486.

PA (UROG-) UROGENESYS INC.

XX

PI Afar DEH, Hubert RS, Raitano AB, Saffran DC, Mitchell SC;
PI Paris M, Jakobovits A;
XX
XX WPI; 2001-367804/38.
DR
XX
XX
PT New STRAP (six transmembrane epithelial antigen of the prostate)
PT proteins, expressed in human cancers, useful for detecting and treating
PT cancer.
XX
XX Example 3C; Page 77; 187pp; English.

PS The present invention relates to human six transmembrane epithelial
CC antigen of the prostate (STRAP) protein. STRAP is a member of cell
CC surface serpentine transmembrane antigens. STRAP gene is used in gene
CC therapy, inhibiting the development or progression of a cancer (eg.
CC prostate, colon, bladder, lung, ovarian and pancreatic) expressing STRAP
CC or inhibiting growth or killing cells expressing STRAP in a patient,
CC comprises administering a vaccine composition to the patient, treating a
CC patient with a cancer that expresses STRAP, or inhibiting growth or
CC killing cells expressing STRAP, comprises administering to the patient a
CC vector encoding single chain monoclonal antibody that comprises the
CC variable domains of the heavy and light chains of the monoclonal antibody
CC that specifically binds to STRAP, such that the vector delivers the
CC single chain monoclonal antibody coding sequence to the cancer cells and
CC the encoded single chain monoclonal antibody is expressed
CC intracellularly. The present sequence is an immunogenic peptide of STRAP-
CC 1. STRAP-1 gene is located on chromosome 7p22.3. This peptide is used to
CC immunise sheep for the generation of sheep polyclonal antibodies towards
CC the amino-terminus of anti-STRAP-1
XX

Sequence 15 AA:

Query Match 100.0%; Score 86; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 8.3e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKMKPRRNLEDDYL 15
Db 1 WKMKPRRNLEDDYL 15

RESULT 3
AAY12305

ID AAY12305 standard; protein; 66 AA.

AC AAY12305;

DT 17-JUN-1999 (first entry)

DE Human 5' EST secreted protein SEQ ID NO:336.

KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.

OS Homo sapiens.

PN WO9906548-A2.

PD 11-FEB-1999.

PF 31-JUL-1998; 98WO-IB001222.

PR 01-AUG-1997; 97US-00905135.

PA (GENST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Iacroleix B;

PI WPI; 1999-153778/13.

DR

DR N-PSDB; AAX41138.
XX
PT New nucleic acids encoding human secreted proteins - obtained from cDNA
PT libraries prepared from e.g. liver, ovary, brain, prostate, kidney, lung,
PT umbilical cord, placenta and colon tissue.
PS Claim 27; Page 677-678; 824pp; English.
XX
CC AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins, and encode the proteins given in AAY1261 to
CC AAY12514, respectively. The proteins given represent the signal peptide
CC and an N-terminal fragment of a secreted protein. The nucleic acid
CC sequences can be used for producing secreted human gene products. They
CC can also be used to develop products for diagnosis and therapy. The
CC proteins obtained may have cytokine activity, cell
CC proliferation/differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptide can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell
CC
SQ Sequence 66 AA:
XX
Query Match 100.0%; Score 86; DB 2; Length 66;
Best Local Similarity 100.0%; Pred. No. 4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 WKMKPRRNLEDDYL 15
DB 14 WKMKPRRNLEDDYL 28
XX
RESULT 4
AAY12304
ID AAY12304 standard; protein; 109 AA.
XX
AC AAY12304;
XX
DT 17-JUN-1999 (first entry)
XX
XX Human 5' EST secreted protein SEQ ID NO:335.
XX
DB Human, secreted protein; EST; expressed sequence tag; diagnosis;
XX forensic; gene therapy; chromosome mapping; signal peptide;
XX upstream regulatory sequence; cytokine activity; cell proliferation;
XX differentiation; haematopoiesis regulation; tissue growth regulation;
XX reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
XX thrombolytic; anti-inflammatory; tumour inhibition.
XX
XX Homo sapiens.
XX
XX WO906548-A2.
XX
XX PD 11-FEB-1999.
XX
XX PF 31-JUL-1998; 98WO-1B001222.
XX
XX PR 01-AUG-1997; 97US-00905135.
XX
XX PA (GEST) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Lacroix B;
XX
XX WPI; 1999-153778/13.
XX
XX DR N-PSDB; AAX41137.
XX
PT New nucleic acids encoding human secreted proteins - obtained from cDNA
PT libraries prepared from e.g. liver, ovary, brain, prostate, kidney, lung,

PT umbilical cord, placenta and colon tissue.
XX
XX Claim 27; Page 677; 824pp; English.
XX
CC AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins, and encode the proteins given in AAY1261 to
CC AAY12514, respectively. The proteins given represent the signal peptide
CC and an N-terminal fragment of a secreted protein. The nucleic acid
CC sequences can be used for producing secreted human gene products. They
CC can also be used to develop products for diagnosis and therapy. The
CC proteins obtained may have cytokine activity, cell
CC proliferation/differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptide can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell
CC
SQ Sequence 109 AA:
XX
Query Match 100.0%; Score 86; DB 2; Length 109;
Best Local Similarity 100.0%; Pred. No. 6.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 WKMKPRRNLEDDYL 15
DB 14 WKMKPRRNLEDDYL 28
XX
RESULT 5
ABU98426
ID ABU98426 standard; protein; 254 AA.
XX
AC ABU98426;
XX
XX 31-JUL-2003 (first entry)
XX
XX STEAP-1 variant 8PID4 v.2 #2.
XX
XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
XX cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX vaccine.
XX
XX Homo sapiens.
XX
XX WO2003022995-A2.
XX
XX PD 20-MAR-2003.
XX
XX PF 06-SEP-2002; 2002MO-US028371.
XX
XX PR 06-SEP-2001; 2001US-0317840P.
XX
XX PR 05-APR-2002; 2002US-0370387P.
XX
XX PA (AGEN-) AGENSYS INC.
XX
XX Paris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
XX
XX WPI; 2003-313240/30.
XX
XX New composition comprising a substance that modulates the status of a
XX STEAP-1-related protein, useful for treating and detecting cancer.
XX
XX Example 53; Page 169-170; 248pp; English.
XX
XX The invention describes a composition comprising a substance that
XX modulates the status of a protein (1) of 340 or 283 amino acids, or of
XX any of the 15 sequences of 259 amino acids, given in the specification,
XX or a molecule that is modulated by the protein, where the status of the

CC cell that expresses the protein is modulated. The compositions, proteins,
CC polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
SQ Sequence 254 AA;

Query Match 100.0%; Score 86; DB 6; Length 254;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKMKPRRLNEEDYL 15
14 WKMKPRRLNEEDYL 28

RESULT 6

ABU98425
ID ABU98425 standard; protein; 254 AA.

AC ABU98425;

DT 31-JUL-2003 (first entry)

DE STEAP-1 variant 8PID4 v.1 #1.

XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX vaccine.

OS Homo sapiens.

PN WC2003022995-A2.

PD 20-MAR-2003.

PF 06-SEP-2002; 2002WO-US028371.

PR 06-SEP-2001; 2001US-0317840P.

PR 05-APR-2002; 2002US-0370387P.

PA (AGEN-) AGENSYS INC.

PI Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;

DR WPI; 2003-313240/30.

PT New composition comprising a substance that modulates the status of a
STEAP-1-related protein, useful for treating and detecting cancer.

PS Example 53; Page 169-170; 248pp; English.

XX The invention describes a composition comprising a substance that
CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
CC any of the 15 sequences of 259 amino acids, given in the specification,
CC or a molecule that is modulated by the protein, where the status of the
CC cell that expresses the protein is modulated. The compositions, proteins,
CC polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
SQ Sequence 254 AA;

Query Match 100.0%; Score 86; DB 6; Length 254;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKMKPRRLNEEDYL 15
14 WKMKPRRLNEEDYL 28

RESULT 7

ABU98429
ID ABU98429 standard; protein; 255 AA.

AC ABU98429;

DT 31-JUL-2003 (first entry)

DE STEAP-1 variant 8PID4 v.3 #2.

XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX vaccine.

OS Homo sapiens.

PN WC2003022995-A2.

PD 20-MAR-2003.

PF 06-SEP-2002; 2002WO-US028371.

PR 06-SEP-2001; 2001US-0317840P.

PR 05-APR-2002; 2002US-0370387P.

PA (AGEN-) AGENSYS INC.

PI Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;

DR WPI; 2003-313240/30.

PT New composition comprising a substance that modulates the status of a
STEAP-1-related protein, useful for treating and detecting cancer.

PS Example 53; Page 172; 248pp; English.

XX The invention describes a composition comprising a substance that
CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
CC any of the 15 sequences of 259 amino acids, given in the specification,
CC or a molecule that is modulated by the protein, where the status of the
CC cell that expresses the protein is modulated. The compositions, proteins,
CC polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
SQ Sequence 255 AA;

Query Match 100.0%; Score 86; DB 6; Length 255;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKMKPRRLNEEDYL 15
14 WKMKPRRLNEEDYL 28

RESULT 8

ABU98428
ID ABU98428 standard; protein; 255 AA.

AC ABU98428;

XX

```

DE 31-JUL-2003 (first entry)
XX
PF STEAP-1 variant 8PID4 v.2 #3.
XX
KM STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
XX cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX vaccine.
XX
OS Homo sapiens.
XX
PN WO2003022995-A2.
XX
PD 20-MAR-2003.
XX
PF 06-SEP-2002; 2002WO-US028371.
XX
PR 06-SEP-2001; 2001US-0317840P.
XX
PR 05-APR-2002; 2002US-0370387P.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
XX
DR WPI; 2003-313240/30.
XX
PT New composition comprising a substance that modulates the status of a
XX STEAP-1-related protein, useful for treating and detecting cancer.
XX
PS Example 53; Page 172; 248pp; English.
XX
CC The invention describes a composition comprising a substance that
CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
CC any of the 15 sequences of 259 amino acids, given in the specification,
CC or a molecule that is modulated by the protein, where the status of the
CC cell that expresses the protein is modulated. The compositions, proteins,
CC polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
XX
SQ Sequence 255 AA;
XX
Query Match 100.0%; Score 86; DB 6; Length 255;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 WKMKPRRLIEDDYL 15
DB 14 WKMKPRRLIEDDYL 28
XX
RESULT 9
ABU98395
ID ABU98395 standard; protein; 258 AA.
XX
AC ABU98395;
XX
DT 31-JUL-2003 (first entry)
XX
DE Novel human gene STEAP-1 variant 13.
XX
KM STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
XX cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX vaccine.
XX
OS Homo sapiens.
XX
PN WO2003022995-A2.
XX
PD 20-MAR-2003.

```

```

XX
PF 06-SEP-2002; 2002WO-US028371.
XX
PR 06-SEP-2001; 2001US-0317840P.
XX
PR 05-APR-2002; 2002US-0370387P.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
XX
DR WPI; 2003-313240/30.
XX
DR N-PSDB; ACD02609.
XX
PT New composition comprising a substance that modulates the status of a
XX STEAP-1-related protein, useful for treating and detecting cancer.
XX
PS Example 2; Fig 2M; 248pp; English.
XX
CC The invention describes a composition comprising a substance that
CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
CC any of the 15 sequences of 259 amino acids, given in the specification,
CC or a molecule that is modulated by the protein, where the status of the
CC cell that expresses the protein is modulated. The compositions, proteins,
CC polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
XX
SQ Sequence 258 AA;
XX
Query Match 100.0%; Score 86; DB 6; Length 258;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 WKMKPRRLIEDDYL 15
DB 14 WKMKPRRLIEDDYL 28
XX
RESULT 10
ABU98389
ID ABU98389 standard; protein; 258 AA.
XX
AC ABU98389;
XX
DT 31-JUL-2003 (first entry)
XX
DE Novel human gene STEAP-1 variant 7.
XX
KM STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
XX cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX vaccine.
XX
OS Homo sapiens.
XX
PN WO2003022995-A2.
XX
PD 20-MAR-2003.
XX
PF 06-SEP-2002; 2002WO-US028371.
XX
PR 06-SEP-2001; 2001US-0317840P.
XX
PR 05-APR-2002; 2002US-0370387P.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
XX
DR WPI; 2003-313240/30.
XX
DR N-PSDB; ACD02603.

```


QY 1 WKMKPRRNLEDDYL 15
 |||||
 DB 14 WKMKPRRNLEDDYL 28

RESULT 13
 ID ABU98387 standard; protein; 258 AA.
 AC ABU98387;
 XX
 XX
 DT 31-JUL-2003 (first entry)
 DE Novel human gene STEAP-1 variant 5.
 XX
 XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
 KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
 KW vaccine.
 XX
 OS Homo sapiens.
 PN WO2003022995-A2.
 XX
 XX 20-MAR-2003.
 PD
 XX
 PF 06-SEP-2002; 2002WO-US028371.
 XX
 PR 06-SEP-2001; 2001US-0317840P.
 PR 05-APR-2002; 2002US-0370387P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 PI Faris M, Ge W, Raitano AB, Challita-Bid PM, Jakobovits A;
 XX
 DR MPI; 2003-313240/30.
 XX
 DR N-PSDB; ACD02601.
 PT New composition comprising a substance that modulates the status of a
 PT STEAP-1-related protein, useful for treating and detecting cancer.
 XX
 PS Example 2; Fig 2B; 248bp; English.
 XX
 CC The invention describes a composition comprising a substance that
 CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
 CC any of the 15 sequences of 259 amino acids, given in the specification,
 CC or a molecule that is modulated by the protein, where the status of the
 CC cell that expresses the protein is modulated. The compositions, proteins,
 CC polynucleotides and methods are useful for treating and detecting cancer.
 CC The STEAP-1-related proteins are useful for generating cancer vaccines.
 CC The polynucleotides are useful as tools for delineating, with greater
 CC precision, cytogenetic abnormalities in the chromosomal region that
 CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
 CC the amino acid sequence of a variant of human six transmembrane
 CC epithelial antigen of the prostate or STEAP-1

SQ Sequence 258 AA;
 Query Match 100.0%; Score 86; DB 6; Length 258;
 Best Local Similarity 100.0%; Pred. No. 1.7e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKMKPRRNLEDDYL 15
 |||||
 DB 14 WKMKPRRNLEDDYL 28

RESULT 14
 ID ABU98393 standard; protein; 258 AA.
 AC ABU98393;
 XX
 XX
 DT 31-JUL-2003 (first entry)

XX Novel human gene STEAP-1 variant 11.
 DE
 XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
 KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
 KW vaccine.
 XX
 OS Homo sapiens.
 PN WO2003022995-A2.
 XX
 XX 20-MAR-2003.
 PD
 XX
 PF 06-SEP-2002; 2002WO-US028371.
 XX
 PR 06-SEP-2001; 2001US-0317840P.
 PR 05-APR-2002; 2002US-0370387P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 PI Faris M, Ge W, Raitano AB, Challita-Bid PM, Jakobovits A;
 XX
 DR MPI; 2003-313240/30.
 XX
 DR N-PSDB; ACD02607.
 PT New composition comprising a substance that modulates the status of a
 PT STEAP-1-related protein, useful for treating and detecting cancer.
 XX
 PS Example 2; Fig 2K; 248bp; English.
 XX
 CC The invention describes a composition comprising a substance that
 CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
 CC any of the 15 sequences of 259 amino acids, given in the specification,
 CC or a molecule that is modulated by the protein, where the status of the
 CC cell that expresses the protein is modulated. The compositions, proteins,
 CC polynucleotides and methods are useful for treating and detecting cancer.
 CC The STEAP-1-related proteins are useful for generating cancer vaccines.
 CC The polynucleotides are useful as tools for delineating, with greater
 CC precision, cytogenetic abnormalities in the chromosomal region that
 CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
 CC the amino acid sequence of a variant of human six transmembrane
 CC epithelial antigen of the prostate or STEAP-1

SQ Sequence 258 AA;
 Query Match 100.0%; Score 86; DB 6; Length 258;
 Best Local Similarity 100.0%; Pred. No. 1.7e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKMKPRRNLEDDYL 15
 |||||
 DB 14 WKMKPRRNLEDDYL 28

RESULT 15
 ID ABU98390 standard; protein; 258 AA.
 AC ABU98390;
 XX
 XX
 DT 31-JUL-2003 (first entry)
 DE Novel human gene STEAP-1 variant 8.
 XX
 XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
 KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
 KW vaccine.
 XX
 OS Homo sapiens.
 PN WO2003022995-A2.
 XX
 XX 20-MAR-2003.

XX 06-SEP-2002; 2002WO-US028371.
 PF
 XX
 PR 06-SEP-2001; 2001US-0317840P.
 PR 05-APR-2002; 2002US-0370387P.
 XX
 PA (AGEN-) AGENSYS INC.

PI Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
 XX
 DR WPI; 2003-313240/30.
 DR N-PSDB; ACD02604.

XX New composition comprising a substance that modulates the status of a
 PT STEAP-1-related protein, useful for treating and detecting cancer.
 XX

PS Example 2; Fig 2H; 248pp; English.

XX The invention describes a composition comprising a substance that
 CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
 CC any of the 15 sequences of 259 amino acids, given in the specification,
 CC or a molecule that is modulated by the protein, where the status of the
 CC cell that expresses the protein is modulated. The compositions, proteins,
 CC polynucleotides and methods are useful for treating and detecting cancer.
 CC The STEAP-1-related proteins are useful for generating cancer vaccines.
 CC The polynucleotides are useful as tools for delineating, with greater
 CC precision, cytogenetic abnormalities in the chromosomal region that
 CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
 CC the amino acid sequence of a variant of human six transmembrane
 CC epithelial antigen of the prostate or STEAP-1

XX Sequence 258 AA;

Query Match 100.0%; Score 86; DB 6; Length 258;
 Best Local Similarity 100.0%; Pred. No. 1.7e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WKMKPRRNLEDDYL 15
 |||||
 Db 14 WKMKPRRNLEDDYL 28

Search completed: July 12, 2004, 14:07:22
 Job time : 33.9651 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:00:58 ; Search time 32.4884 Seconds
(without alignments)
213.658 Million cell updates/sec

Title: US-10-010-667a-2_COPY_92_113
Perfect score: 116
Sequence: 1 REVHPLATSHOOYFYKIPILV 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL 25: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_virus: *
16: sp_bacteriaph: *
17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	108	93.1	338	6	Q9GL50	Q9GL50	sus scrofa
2	94	81.0	339	11	Q924J9	Q924J9	mus musculus
3	94	81.0	339	11	Q924Z2	Q924Z2	mus musculus
4	94	81.0	339	11	Q9CNR7	Q9CNR7	mus musculus
5	68	58.6	489	11	Q8BWB6	Q8BWB6	mus musculus
6	68	58.6	490	4	Q8NFW2	Q8NFW2	homo sapien
7	68	58.6	490	4	Q8IDU7	Q8IDU7	homo sapien
8	50	43.1	172	16	Q8NZR1	Q8NZR1	streptococc
9	43	42.2	764	16	Q7UR83	Q7UR83	rhodospirill
10	48	41.4	172	16	Q938U7	Q938U7	streptococc
11	48	41.4	388	16	Q9A0P6	Q9A0P6	streptococc
12	48	41.4	456	4	Q8TDP3	Q8TDP3	homo sapien
13	48	41.4	487	4	Q8TF03	Q8TF03	homo sapien
14	48	41.4	487	4	Q8NFW6	Q8NFW6	homo sapien
15	48	41.4	488	4	Q9NVB5	Q9NVB5	homo sapien
16	48	41.4	488	4	Q86SR6	Q86SR6	homo sapien

17	48	41.4	498	4	Q7Z389	Q7Z389	homo sapien
18	48	41.4	676	13	Q7ZV52	Q7ZV52	brachydanio
19	48	41.4	929	12	Q9QAP4	Q9QAP4	cervid herp
20	47	40.5	172	16	Q8ES61	Q8ES61	oceanobacti
21	47	40.5	233	10	Q7XUP5	Q7XUP5	arabidopsis
22	47	40.5	427	16	Q8R872	Q8R872	thermoanaer
23	47	40.5	429	16	Q8KES8	Q8KES8	chlorobium
24	47	40.5	431	17	Q89V23	Q89V23	bradyrhizob
25	47	40.5	431	17	Q9HMR3	Q9HMR3	halobacteri
26	47	40.5	434	16	Q8F4H9	Q8F4H9	leptospira
27	47	40.5	438	16	Q7TVW8	Q7TVW8	synchococc
28	47	40.5	439	16	Q7UG05	Q7UG05	rhodospirill
29	47	40.5	442	16	Q7VCE0	Q7VCE0	prochloroco
30	47	40.5	442	16	Q7TU77	Q7TU77	prochloroco
31	47	40.5	442	16	Q7TUC3	Q7TUC3	prochloroco
32	47	40.5	565	3	Q04600	Q04600	saccharomyc
33	47	40.5	568	4	Q96MU3	Q96MU3	homo sapien
34	47	40.5	675	3	Q9UUE5	Q9UUE5	schizosacch
35	47	40.5	688	10	Q7XUG5	Q7XUG5	oryza sativ
36	47	40.5	749	4	Q8NB24	Q8NB24	homo sapien
37	47	40.5	763	10	Q9LMT2	Q9LMT2	oryza sativ
38	47	40.5	763	10	Q9LGD6	Q9LGD6	oryza sativ
39	47	40.5	763	10	Q9FV78	Q9FV78	oryza sativ
40	47	40.5	806	10	Q9LMM2	Q9LMM2	oryza sativ
41	47	40.5	912	12	Q8OV01	Q8OV01	buffalo herp
42	47	40.5	947	12	Q9QAP7	Q9QAP7	bovine herp
43	47	40.5	1452	4	Q8C0F0	Q8C0F0	homo sapien
44	46.5	40.1	437	2	Q93TY7	Q93TY7	shigella fl
45	46.5	40.1	657	16	Q8NZ96	Q8NZ96	streptococc

ALIGNMENTS

RESULT 1
ID Q9GL50 PRELIMINARY; PRT; 338 AA.
AC Q9GL50;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Six transmembrane endothelial antigen of PAEC.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Nagasaka T., Bouliday G., Coupel S., Coulon F., Tesson L.,
RA Heslan J.-M., Soullion J.-P., Charreau B.;
RT "Differential gene expression in endothelial cells during TNF-alpha-
RT and IPS-mediated activation.";
RI Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF19659; AAG33868.1; -
DR GO: GO:0016021; C:Integral to membrane; IEA.
KW Transmembrane.
SQ
SEQUENCE 338 AA: 39918 MW; ED490E8E067A32B CRC64;
Query Match 93.1%; Score 108; DB 6; Length 338;
Best Local Similarity 90.9%; Pred. No. 8.1e-10;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 REVHPLATSHOOYFYKIPILV 22
DB 91 REVHPLATSHOOYFYKIPILV 112
RESULT 2
ID Q924J9 PRELIMINARY; PRT; 339 AA.
AC Q924J9;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

```

DT 01-JUN-2003 (TREMELREL. 24, Last annotation update)
DE Six-transmembrane epithelial antigen of the prostate.
GN STRAP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=21371909; PubMed=11479226;
RA Yang D., Holt G.E., Velders M.P., Kwon E.D., Kast W.M.;
RT "Murine six-transmembrane epithelial antigen of the prostate, prostate
RT stem cell antigen, and prostate-specific membrane antigen: prostate-
RT specific cell-surface antigens highly expressed in prostate cancer of
RT transgenic adenocarcinoma mouse prostate mice.";
RL Cancer Res. 61:5857-5860(2001).
DR EMBL; AF297098; AAK83126.1; -.
DR MGD; MGI:1917608; Steap.
DR GO; GO:0016021; C:integral to membrane; IEA.
KW Transmembrane.
SQ SEQUENCE 339 AA; 39318 MW; 4B26A71FF559E84F CRC64;

Query Match 81.0%; Score 94; DB 11; Length 339;
Best Local Similarity 77.3%; Pred. No. 1.9e-07;
Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 REV1HPLATSHOQFYKIPILV 22
Db 92 RETIYPLVTSREQFYKIPILV 113

RESULT 3
Q924Z2 PRELIMINARY; PRT; 339 AA.
ID 0924Z2;
AC Q924Z2;
DT 01-DEC-2001 (TREMELREL. 19, Created)
DT 01-DEC-2001 (TREMELREL. 19, Last sequence update)
DT 01-JUN-2003 (TREMELREL. 24, Last annotation update)
DE Dudulin.
GN STEAP OR 1010001D01RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Serra V., Nantvet P., Lambdin D., Vaubourgoille M., Kellermann O.,
RA Loric S.;
RT "Prostate and non-prostate expression of dudulin, the mouse ortholog
RT of human STEAP.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY029584; AAK50537.1; -.
DR MGD; MGI:1917608; Steap.
SQ SEQUENCE 339 AA; 39109 MW; 32A2C29F2E333BD0 CRC64;

Query Match 81.0%; Score 94; DB 11; Length 339;
Best Local Similarity 77.3%; Pred. No. 1.9e-07;
Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 REV1HPLATSHOQFYKIPILV 22
Db 92 RETIYPLVTSREQFYKIPILV 113

RESULT 4
Q9CWR7 PRELIMINARY; PRT; 339 AA.
AC Q9CWR7;
DT 01-JUN-2001 (TREMELREL. 17, Created)
DT 01-JUN-2001 (TREMELREL. 17, Last sequence update)
DT 01-OCT-2002 (TREMELREL. 22, Last annotation update)
DE 2410007B19RIK protein.

```

```

GN STRAP OR 2410007B19RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Embryonic stem cells;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Kawai T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batcalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima Y., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010437; BAB26938.1; -.
DR MGD; MGI:1917608; Steap.
SQ SEQUENCE 339 AA; 39264 MW; 3F7AB9C7520F0968 CRC64;

Query Match 81.0%; Score 94; DB 11; Length 339;
Best Local Similarity 77.3%; Pred. No. 1.9e-07;
Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 REV1HPLATSHOQFYKIPILV 22
Db 92 RETIYPLVTSREQFYKIPILV 113

RESULT 5
Q8BWB6 PRELIMINARY; PRT; 489 AA.
ID Q8BWB6;
AC Q8BWB6;
DT 01-MAR-2003 (TREMELREL. 23, Created)
DT 01-MAR-2003 (TREMELREL. 23, Last sequence update)
DT 01-JUN-2003 (TREMELREL. 24, Last annotation update)
DE Weakly stimilar to tumor suppressor pRbE.
GN 4921538B17RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK052981; BAC35230.1; -.
DR MGD; MGI:1921301; 4921538B17RIK.
DR GO; GO:0006118; P:electon transport; IEA.
DR InterPro: IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420_oxidored; I.
SQ SEQUENCE 489 AA; 55760 MW; 98CD63D59DDDF24C CRC64;

Query Match 58.6%; Score 68; DB 11; Length 489;
Best Local Similarity 65.0%; Pred. No. 0.007;

```

Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 REVHPLATSHOQFYKIP1 20
|:|||||:|
Db 233 RDVHPYARNQSDFKIP1 251

RESULT 6

Q8NFT2 PRELIMINARY; PRT; 490 AA.

AC Q8NFT2; (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, last sequence update)
DT 01-OCT-2002 (TREMBlrel. 24, last annotation update)
DE Six-transmembrane epithelial antigen of prostate 2.
GN STEAP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Porikha K.P., Helenius M.A., Viaskorpi T.;
RT "Cloning and characterization of a novel six-transmembrane protein
STEAP2, down-regulated in androgen-independent prostate cancer."
RT Submitted (DEC-2001) to the EMBL/GenBank/DBD databases.
DR EMBL; AA45138; AAN04080.1; -
DR GeneW; HGNC:17885; STEAP2.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420_Oxidored; 1.
KW Transmembrane.

SEQUENCE 490 AA; 55961 MW; P2E9C30CDACCF81 CRC64;

Query Match 58.6%; Score 68; DB 4; Length 490;

Best Local Similarity 65.0%; Pred. No. 0.0071; 5; Indels 0; Gaps 0;

Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 1 REVHPLATSHOQFYKIP1 20
|:|||||:|
Db 233 RDVHPYARNQSDFKIP1 252

RESULT 7

Q8IUE7 PRELIMINARY; PRT; 490 AA.

AC Q8IUE7; (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, last sequence update)
DT 01-MAR-2003 (TREMBlrel. 24, last annotation update)
DE STAMP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Korkmaz K.S., Elbi C.C., Korkmaz C.G., Loda M., Hager G.L.,
RA Saatioglu F.;
RT "Molecular cloning and characterization of STAMP1, a highly prostate
specific six-trans-membrane protein that is overexpressed in prostate
cancer";
RT J. Biol. Chem. 277:36689-36696(2002).

DR EMBL; AY008445; AAG32149.1; -
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420_Oxidored; 1.
DR STRAIN=1;
SEQUENCE 490 AA; 56027 MW; D1971A8480F7E51 CRC64;

Query Match 58.6%; Score 68; DB 4; Length 490;

Best Local Similarity 65.0%; Pred. No. 0.0071; 5; Indels 0; Gaps 0;

QY 1 REVHPLATSHOQFYKIP1 20
|:|||||:|
Db 233 RDVHPYARNQSDFKIP1 252

RESULT 8

Q8NZR1 PRELIMINARY; PRT; 172 AA.

AC Q8NZR1; (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, last annotation update)
DE Hypothetical phage protein spvM18_1784.
GN spvM18_1784.
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_Taxid=186103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbhan K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
group A Streptococcus strains associated with acute rheumatic fever
outbreaks";
RT Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).

DR EMBL; AE010086; AAL98305.1; -
DR GO; GO:0005727; C:extrachromosomal circular DNA; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008152; P:metabolism; IEA.

DR InterPro; IPR005811; CoA_ligase.
DR InterPro; IPR003115; ParBc.
DR Pfam; PF02195; ParBc; 1.
KW Complete proteome; Hypothetical protein.

SEQUENCE 172 AA; 19672 MW; BE81A8BD93BB7EA6 CRC64;

Query Match 43.1%; Score 50; DB 16; Length 172;

Best Local Similarity 38.1%; Pred. No. 2.5; 5; Indels 0; Gaps 0;

Matches 8; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
QY 2 EVHPLATSHOQFYKIP1V 22
|:|||||:|
Db 233 EAVEPYAESIKFEGFKIPV 43

RESULT 9

Q7UL83 PRELIMINARY; PRT; 764 AA.

AC Q7UL83; (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, last annotation update)
DE Inter-alpha-trypsin inhibitor family heavy chain-related
protein-hypothetical secreted or membrane-associated protein
containing WFPA domain.
GN RB9664.
OS Rhodospirillum rubrum.

OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_Taxid=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
MEDLINE=22735913; PubMed=12835416;

RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
 RA Ludwig W., Gade D., Beck A., Borzym K., Hellmann K., Rabus R.,
 RA Schlegner H., Mann R., Reinhardt R.,
 RT "Complete genome sequence of the marine planctomycete *Pirellula* sp.
 RT strain 1.",
 RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).
 DR EMBL; BX294150; CAD76394.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 764 AA; 82826 MW; 5A77412DB5AE3311 CRC64;

Query Match 42.2%; Score 49; DB 16; Length 764;
 Best Local Similarity 33.3%; Pred. No. 19;
 Matches 7; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 2 EVIHPATSHQOYFYKPIIV 22
 Db 183 EYVHPLEIDEDRYVFRPMV 203

RESULT 10

Q938L7 PRELIMINARY; PRT; 172 AA.

AC Q938L7, 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical phage associated protein SplyM3_1236 (Phage
 DE associated).
 GN SPLYM3_1236 OR SPS0625.
 OS Streptococcus pyogenes, and
 OC Streptococcus pyogenes (serotype M3).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314, 198466;
 RN
 RP SEQUENCE FROM N.A.
 RC SPECIES=S. pyogenes; STRAIN=NIH1;
 RA Ikebe T., Wada A., Inagaki Y., Sugama K., Tanaka D., Suzuki R.,
 RA Katsukawa C., Fujiwara Y., Abe Y., Watanabe H.;
 RT "Complete sequence of temperate phage PhiNIH1.1";
 RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE FROM N.A.
 RC SPECIES=S. pyogenes; STRAIN=MGAS315 / Serotype M3;
 RX MEDLINE=22133808; PubMed=12122206;
 RA Beres S.B., Sylva G.L., Barbhan K.D., Lei B., Hoff J.S.,
 RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
 RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
 RA Schlievert P.M., Musser J.M.;
 RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
 RT phage-encoded toxins, the high-virulence phenotype, and clone
 RT emergence";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083 (2002).
 RN
 RP SEQUENCE FROM N.A.
 RC SPECIES=S. pyogenes; STRAIN=SSI-1 / Serotype M3;
 RA Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
 RA Yamazaki K., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,
 RA Hayashi H., Hamada S.;
 RT "The genome of invasive Streptococcus pyogenes; a comparative analysis
 RT of S. pyogenes SSI-1, SF370 and MGAS8232.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY050245; AL15065.1; -;
 DR EMBL; AE014160; AA079843.1; -;
 DR EMBL; AB005143; BAC63720.1; -;
 DR GO; GO:0005727; C:extrachromosomal circular DNA; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR InterPro; IPR003115; ParBc.
 DR Pfam; PF02195; ParBc; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 172 AA; 19484 MW; ADB1E6565EDAE4C CRC64;

Query Match 41.4%; Score 48; DB 16; Length 172;

Best Local Similarity 38.1%; Pred. No. 5.4;
 Matches 8; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 2 EVIHPATSHQOYFYKPIIV 22
 Db 23 EAVGPVASEIKERGFVPIV 43

RESULT 11

Q9A0P6 PRELIMINARY; PRT; 388 AA.

AC Q9A0P6, 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical phage associated protein.
 GN SPLY0679.
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
 RX MEDLINE=21192684; PubMed=11296296;
 RA Perretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
 RA Primeaux C., Szate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
 RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
 RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663 (2001).
 RL
 RM EMBL; AB006521; AKK33638.1; -;
 DR GO; GO:0005727; C:extrachromosomal circular DNA; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
 DR GO; GO:0006306; P:DNA methylation; IEA.
 DR InterPro; IPR002295; D21N6_mtfase.
 DR InterPro; IPR001091; Met_Tfns_CN4.
 DR InterPro; IPR002941; N6_M4_Mcase.
 DR InterPro; IPR002052; N6_M4_Mcase.
 DR InterPro; IPR003115; ParBc.
 DR Pfam; PF01555; N6_N4_Mcase; 1.
 DR Pfam; PF02195; ParBc; 1.
 DR PRINTS; PR00506; D21N6TRANSFER.
 DR PRINTS; PR00508; S21N4MTFRASE.
 DR SMART; SM00470; ParB; 1.
 DR PROSITE; PS00092; N6_MTASE; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 388 AA; 43690 MW; 12587B620606BB7 CRC64;

Query Match 41.4%; Score 48; DB 16; Length 388;
 Best Local Similarity 38.1%; Pred. No. 13;
 Matches 8; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 2 EVIHPATSHQOYFYKPIIV 22
 Db 23 EAVGPVASEIKERGFVPIV 43

RESULT 12

Q8TDP3 PRELIMINARY; PRT; 456 AA.

AC Q8TDP3, 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE PHYR II.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN
 RP SEQUENCE FROM N.A.

RA Wang C., Alley J.A., Steiner M.S.;
RT "Second human member of pHyde family, Human pHyde II."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF262322; AAM08128.1; -
DR GO:GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; I.
DR PROSITE; PS00290; IG_MHC; I.
SQ SEQUENCE 456 AA; 50430 MW; C5F7C7008D55251E CRC64;

Query Match 41.4%; Score 48; DB 4; Length 456;
Best Local Similarity 36.4%; Pred. No. 16;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFKIPILV 22
DB 233 RDVLPYVQESQNKFFKLPVSV 254

RESULT 13
ID Q8TF03 PRELIMINARY; PRT; 487 AA.
AC Q8TF03;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
DE Tumor suppressor pHyde.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Steiner M.S., Alley J.A., Wang C.;
RT "A novel prostate-derived tumor suppressor."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF238864; AAL78206.1; -
DR GO:GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; I.
DR PROSITE; PS00290; IG_MHC; I.
SQ SEQUENCE 487 AA; 54402 MW; AFR16053590E6F68 CRC64;

Query Match 41.4%; Score 48; DB 4; Length 487;
Best Local Similarity 36.4%; Pred. No. 17;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFKIPILV 22
DB 233 RDVLPYVQESQNKFFKLPVSV 254

RESULT 14
ID Q8NEM6 PRELIMINARY; PRT; 487 AA.
AC Q8NEM6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
DE pHyde.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lu Y., Beheshti B., Squire J.A., Yang X.J.;
RT "Characterization of a novel apoptosis-inducing gene, hpHyde, that
inhibits prostate cancer cell growth."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY082673; AAM45136.1; -

DR GO:GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; I.
DR PROSITE; PS00290; IG_MHC; I.
SQ SEQUENCE 487 AA; 54501 MW; E0EP9585C81EDF CRC64;

Query Match 41.4%; Score 48; DB 4; Length 487;
Best Local Similarity 36.4%; Pred. No. 17;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFKIPILV 22
DB 233 RDVLPYVQESQNKFFKLPVSV 254

RESULT 15
ID Q9NVB5 PRELIMINARY; PRT; 488 AA.
AC Q9NVB5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
DE Hypothetical protein FLJ10829 (Dudulin 2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuno Y., Sasaki N.;
RT "NEPO human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Serru V., Manivet P., Lenoir C., Eschwege F., Lambdin D.,
RA Vauthourdolle M., Kellermann O., Loric S.;
RT "Dudulin 2, a new tumor antigen expressed in various human tumors."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK001691; BAA1839.1; -
DR EMBL; AY029585; AAK50538.1; -
DR GO:GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; I.
DR PROSITE; PS00290; IG_MHC; I.
KW Hypothetical protein.
SQ SEQUENCE 488 AA; 54616 MW; BC0BCA48335FAD6 CRC64;

Query Match 41.4%; Score 48; DB 4; Length 488;
Best Local Similarity 36.4%; Pred. No. 17;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFKIPILV 22
DB 233 RDVLPYVQESQNKFFKLPVSV 254

Search completed: July 12, 2004, 14:10:24
Job time : 33.4884 secs

This Page Blank (uspto)

Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 13:52:02 ; Search time 6.90698 Seconds

(without alignments)
165.853 Million cell updates/sec

Title: US-10-010-667a-2_COPY_92_113
Perfect score: 116
Sequence: 1 REVHPLATSHQYFYKPIIV 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	100.0	339	1	STEPA HUMAN
2	52.5	45.3	465	1	CIXG HAEIN
3	49	42.2	177	1	MSRA_HAUNI
4	47.5	40.9	320	1	CYP_FORPU
5	46	39.7	485	1	STPA_ARATH
6	46	39.7	701	1	GLGE_MYCTU
7	46	38.8	1839	1	CYPA_SACKL
8	45	38.7	304	1	WR13_ARATH
9	45	38.8	2549	1	FRAP_MOUSE
10	44.5	38.4	137	1	SPMI_PIG
11	44.5	38.4	308	1	PRB_METAC
12	44.5	38.4	1176	1	RPE2_DROME
13	44	37.9	811	1	RPX1_YEAST
14	44	37.9	928	1	VGLB_HSVBP
15	44	37.9	932	1	VGLB_HSVBP
16	43.5	37.5	326	1	YHNS_ECOLI
17	43	37.1	337	1	DHAS_VITCH
18	43	37.1	373	1	EXT1_ARATH
19	43	37.1	385	1	GBAS_CAEEL
20	42.5	36.6	299	1	PYRB_ARCFU
21	42	36.2	216	1	MSRA_XYLFU
22	42	36.2	216	1	MSRA_XYLFU
23	42	36.2	316	1	DHAS_VITAM
24	42	36.2	580	1	PIGO_MOUSE
25	42	36.2	1437	1	DPO3_BACSV
26	42	36.2	2549	1	FRAP_HUMAN
27	42	36.2	2549	1	FRAP_HUMAN
28	41.5	35.8	116	1	AON3_PIG
29	41.5	35.8	154	1	MTE_ALUMI
30	41	35.3	194	1	COAE_AGRIS
31	41	35.3	196	1	COAE_HELPJ
32	41	35.3	196	1	COAE_HELPJ
33	41	35.3	216	1	MSRA_XANCP

34	41	35.3	575	1	METW_YEAST
35	41	35.3	576	1	SBP_CAEEL
36	41	35.3	623	1	Y014_MYCPN
37	41	35.3	635	1	XYNB_BUTEI
38	40.5	34.9	309	1	PYRB_METWA
39	40.5	34.9	350	1	XYLD_AGRIS
40	40.5	34.9	433	1	PEPD_SALDU
41	40.5	34.9	492	1	SYNM_YEAST
42	40	34.5	97	1	GATC_PROMP
43	40	34.5	115	1	GUAN_RAT
44	40	34.5	246	1	YNU0_YEAST
45	40	34.5	317	1	PIC_LISMO

ALIGNMENTS

```
RESULT 1
STEPA_HUMAN          STANDARD;          PRT;          339 AA.
AC      Q90HE8; Q95034;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      15-MAR-2004 (Rel. 43, Last annotation update)
DE      Six transmembrane epithelial antigen of prostate.
GN      STEAP OR STEAP1.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_TaxId=9606;
[1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=20056277; PubMed=10586738;
RA      Hubert R.S., Vivanco I., Chen E., Rastegar S., Jeong K.,
RA      Mitchell S.C., Madraswala R., Zhou Y., Kuo J., Rattano A.B.,
RA      Jakobovits A., Saffran D.C., Afar D.B.H.;
RT      "STEAP: a prostate-specific cell-surface antigen highly expressed in
RT      human prostate tumors."
RL      Proc. Natl. Acad. Sci. U.S.A. 96:14523-14528 (1999).
RN      [2]
RN      SEQUENCE FROM N.A.
RP      Abu-Threideh J., Stoneking T., Langston Y., Maupin R.;
RL      Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.
RN      [3]
RN      SEQUENCE FROM N.A.
RC      TISSUE=Skin;
RX      MEDLINE=22388257; PubMed=12477932;
RA      Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Scheefter C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,
RA      Dichtlein L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA      Brownstein M.J., Ustin T.B., Toshitsuki S., Carninci R.D., Mullany S.J.,
RA      Raha S.S., Loguellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA      Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gumatratte P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA      Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA      Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA      Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA      Schercher A., Schein J.E., Jones S.J.M., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length
RT      human and mouse cDNA sequences."
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[1] SUBCELLULAR LOCATION: Integral membrane protein (Potential).
[1] TISSUE SPECIFICITY: Highly expressed in prostate tumors.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
```

use, by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AF186249; AAF17479.1; -.
DR EMBL; AC005053; AAC79150.1; ALT_INIT.
DR EMBL; AC004969; AAD15620.2; -.
DR EMBL; BC011802; AAH1802.1; -.
DR Genew; HGNC:11378; STEAP.
DR MIM; 604415; -.
DR GO; GO:000587; C:integral to plasma membrane; TAS.
DR GO; GO:0005911; C:intracellular junction; TAS.
DR GO; GO:0015267; F:channel/pore class transporter activity; TAS.
KW Transmembrane; Antigen.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 119 139 POTENTIAL.
FT TRANSMEM 164 184 POTENTIAL.
FT TRANSMEM 218 238 POTENTIAL.
FT TRANSMEM 258 278 POTENTIAL.
FT TRANSMEM 291 311 POTENTIAL.
SQ SEQUENCE 339 AA; 39851 MW; 55443A170C870387 CRC64;

Query Match 100.0%; Score 116; DB 1; Length 339;
Best Local Similarity 100.0%; Pred. No. 3.5e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 REVHPLATSHOQFYKIPILV 22
Db 92 REVHPLATSHOQFYKIPILV 113

ID _CITX_HAEIN STANDARD; PRT; 465 AA.
AC P4458;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE CItXg protein [includes: Apo-citrate lyase phosphoribosyl-dephospho-CoA transferase (EC 2.7.7.61) (Holo-citrate lyase synthase) (Apo-ACP nucleosidyltransferase)]; 2-(5''-triphosphoribosyl)-3''-dephospho-CoA synthase (EC 2.7.8.25) (2-(5''-triphosphoribosyl)-3''-dephospho-CoA synthase) (EC 2.7.8.25) (2-(5''-triphosphoribosyl)-3''-dephospho-CoA synthase).
GN CItXG OR CItX OR HI0021.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales; Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F., Kertész A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M., McKenny K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D., Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M., Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D., Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C., Fine L.D., Fritchman J.L., Fuhmann J.L., Georgagen N.S.M., Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O., Venter J.C.;
RA "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.";
RT Science 269:496-512(1995).
RL
CC -I- FUNCTION: Bifunctional enzyme that catalyzes formation of 2-(5''-triphosphoribosyl)-3''-dephospho-CoA, and then the transfer of this prosthetic group precursor to the apo-acyl carrier protein (gamma chain) of the citrate lyase to yield the holo-acyl carrier protein (by similarity).
CC -I- CATALYTIC ACTIVITY: 2-(5''-triphosphoribosyl)-3''-dephospho-CoA + apo-citrate lyase = holo-citrate lyase + diphosphate.
CC -I- CATALYTIC ACTIVITY: ATP + 3-dephospho-CoA = 2''-(5''-

triphosphoribosyl)-3''-dephospho-CoA + adenine.
CC -I- SIMILARITY: In the N-terminal section; belongs to the citX family.
CC -I- SIMILARITY: In the C-terminal section; belongs to the citX/mcCB family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; U32687; AAC21699.1; ALT_INIT.
DR TIGR; HI0021; -.
DR HAMAP; MF_00397; fused; 1.
DR HAMAP; MF_00398; fused; 1.
DR InterPro; IPR002736; CItX.
DR InterPro; IPR005551; CItX.
DR Pfam; PF01874; CItX; 1.
DR Pfam; PF03802; CItX; 1.
DR Transferase; Nucleosidyltransferase, Multifunctional enzyme;
KW Complete proteome.
FT DOMAIN 1 182
FT DOMAIN 183 465
FT DOMAIN 465 51834 MW; 4467221CDB58A96 CRC64;
SQ SEQUENCE 465 AA; 51834 MW; 4467221CDB58A96 CRC64;

Query Match 45.3%; Score 52.5; DB 1; Length 465;
Best Local Similarity 50.0%; Pred. No. 0.84;
Matches 10; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

Qy 1 REVHPLATSHOQFYKIPILV 20
Db 83 KEIRPLETGHAEVYV-LPI 101

ID MSRA_HAINT STANDARD; PRT; 177 AA.
AC Q9HGO0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Peptide methionine sulfoxide reductase msra (EC 1.8.4.6) (Protein-methionine-S-oxide reductase) (Peptide Met (O) reductase).
GN MSRA OR VNG1180G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteriia; Halobacteriales; Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng M.V., Kennedy S.P., Mahlras G.G., Bergquist B., Pan M., Shukla H.D., Lasky S.R., Baliga N.S., Thorson V., Shrooga J., Swartzell S., Weir D., Hall J., Dahl T.A., Weir R., Goo Y.A., Letshauser B., Keller K., Cruz R., Danson M.J., Kough D.W., Maddocks D.G., Jablonski P.B., Krebs M.P., Angevine C.M., Dale H., Isebnarger T.A., Peck R.F., Pohlshroder M., Spudich J.L., Jung K.-H., Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.F., Omer A.D., Ebbhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RA "Genome sequence of Halobacterium species NRC-1";
RT Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RL
CC -I- FUNCTION: Has an important function as a repair enzyme for proteins that have been inactivated by oxidation. Catalyzes the reversible oxidation-reduction of methionine sulfoxide in proteins to methionine (by similarity).
CC -I- CATALYTIC ACTIVITY: Protein L-methionine + oxidized thioredoxin = protein L-methionine S-oxide + reduced thioredoxin.
CC -I- SIMILARITY: Belongs to the msra Met sulfoxide reductase family.
CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AE005045; AAG19555.1; -.
DR PIR: G84273; G84273.
DR HSSP: P54149; 1FWA.
DR HAMAP: ME_01401; -; 1.
DR InterPro: IPR002569; PMSR.
DR Pfam: PF01625; PMSR; 1.
DR ProDom: PD003489; PMSR; 1.
DR TIGRFAMs: TIGR00401; msrA; 1.
DR Oxidoreductase; Complete proteome.
FM ACT SITE 12 BY SIMILARITY.
FT SEQUENCE 177 AA; 19202 MW; 337FBD2E1CFDD8EC CRC64;
SQ

Query Match 42.2%; Score 49; DB 1; Length 177;
Best Local Similarity 50.0%; Pred. No. 1.1;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 EVINPLATSHQGYFYKIP 19
DB 135 ETFYPAEKHKYFKENP 152

RESULT 4
CYF_PORPU STANDARD; PRT; 320 AA.
AC P51265;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Apocytochrome f precursor.
GN PERA.
OS Porphyra purpurea.
OG Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
OX NCBI_Taxid=2787;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Avonport;
RA Reich M.E., Munholland J.;
RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast
genome.";
RL Plant Mol. Biol. Rep. 13:333-335(1995).
CC -!- FUNCTION: Translocates protons across the thylakoid membrane and
CC transfers electrons from photosystem II to photosystem I. It
CC receives electrons from the Rieske iron-sulfur protein and passes
CC them to plastocyanin.
CC -!- SUBUNIT: Interacts with plastocyanin and Rieske iron-sulfur
CC protein (By similarity).
CC -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane. Membrane-
CC anchored (By similarity).
CC -!- SIMILARITY: Belongs to the cytochrome c family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U38804; AAC08151.1; -.
DR PIR: S73186; S73186.
DR HSSP: P36438; 1HCZ.
DR HAMAP: MF_00610; -; 1.
DR InterPro: IPR002325; Apocyt_F.
DR InterPro: IPR000345; CytC_heme_BS.

DR Pfam: PF01333; Apocytochr F C; 1.
DR PRINTS: PR00610; CYTOCHROME_F.
DR PROSITE: PS00190; CYTOCHROME_C; 1.
KM Electron transport; Heme; Chloroplast; Thylakoid; Photosynthesis;
KM Photosystem I; Photosystem II; Transmembrane; Signal.
FT SIGNAL 1 36
FT CHAIN 37 320
FT TRANSMEM 286 305
FT METAL 37 37
FT BINDING 57 57
FT BINDING 60 60
FT METAL 61 61
SQ SEQUENCE 320 AA; 35012 MW; C9B103F0453369C7 CRC64;
QY 1 REVHPL-----ATSHQGYFYKIPLV 22
DB 160 REIIFPLSPDPKXKQAHFFKYPYIV 186

Query Match 40.9%; Score 47.5; DB 1; Length 320;
Best Local Similarity 40.7%; Pred. No. 3.6;
Matches 11; Conservative 4; Mismatches 7; Indels 5; Gaps 1;

RESULT 5
SYFA_ARATH STANDARD; PRT; 485 AA.
ID SYFA_ARATH
AC Q9T034; O8LPR1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Probable phenylalanine-tRNA synthetase alpha chain (EC 6.1.1.20)
DE (Phenylalanine-tRNA ligase alpha chain) (PheRS).
GN AY439280 OR T22F8.180.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_Taxid=3702;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schneller C., Wandutt R., Murphy G., Voickaert G.,
RA Pohl T., Duesterhoelt A., Stiekema W., Ertian X.-D., Terry N.,
RA Harris B., Ansoorge W., Brandt P., Griwall L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermayer B., Macher R., Mueller M.,
RA Kreis M., Deleney M., Puigdomenech P., Watson M., Schmidtnein T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hohnsels U., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Grymoprez B., Chuang Y.-Y., Vandenbussche F.,
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA Holzner E., Brandt A., Peters S., van Staveren M., Dirke W.,
RA Woolman P., Klein Lankhorst R., Rose M., Hauf J., Koelter P.,
RA Bernisier S., Hempel S., Feldpausch M., Jambert S., Van den Daele H.,
RA De Keyser A., Buysaert C., Gielen J., Villartiel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McLeay K., Mayes R.,
RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann T.-H.,
RA Borkova D., Blocher H., Scharfe M., Grimm M., Mueller-Auer S.,
RA Dose S., Fuchs M., Fartman B., Grandenath K., Dauber D., Heitzl A.,
RA Neumann S., Argiriou A., Vitale D., Lignot R., Piravatti E.,
RA Messner O., Outley F., Clabaud G., Muendlein A., Felber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Chedfor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Barques M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Jacou D., Uesse T.,
RA Heijman L., Schwarz S., Scholler P., Heber S., Frances P., Bietke C.,
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,

DR InterPro; IPR006047; Alpha_amyl_cat.
 DR Pfam; PF00128; alpha-amylase; 1.
 DR SMART; SM00642; Amy; 1.
 DR Hydrolase; Glycosidase; Complete proteome.
 SQ SEQUENCE 701 AA; 78639 MW; EE637BA1DA7D694F CRC64;

Query Match 39.7%; Score 46; DB 1; Length 701;
 Best Local Similarity 46.7%; Pred. No. 15;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 HPLATSHQCYFKIP 19
 DB 353 HPWAHHRQWTELP 367

RESULT 7
 ID CYAA_SACKL STANDARD; PRT; 1839 AA.
 AC P23466;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Adenylylase cyclase (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylyl cyclase).
 GN CYRL.
 OS Saccharomyces kluyveri (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4934;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91323718; PubMed=1864503;
 RA Young D., O'Neill K., Broek D., Wiegler M.;
 RT "The adenylyl cyclase-encoding gene from Saccharomyces kluyveri.";
 RL Gene 102:123-132(1991).
 CC -!- FUNCTION: Plays essential roles in regulation of cellular metabolism by catalyzing the synthesis of a second messenger, cAMP.
 CC -!- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
 CC -!- COFACTOR: Binds 1 magnesium ion per subunit (BY similarity).
 CC -!- SIMILARITY: Belongs to the adenylyl cyclase class-3 family.
 CC -!- SIMILARITY: Contains 21 leucine-rich (LRR) repeats.
 CC -!- SIMILARITY: Contains 1 PP2C-like domain.
 CC -!- SIMILARITY: Contains 1 Ras-associating domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X56042; CA39513.1; -.
 DR PIR; J01145; OYBYK.
 DR InterPro; IPR001054; G_cyclase.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR003591; LRR_Typ.
 DR InterPro; IPR001932; PP2C-like.
 DR InterPro; IPR000159; RA_domain.
 DR Pfam; PF00211; guanylate_cyc; 1.
 DR Pfam; PF00560; LRR; 15.
 DR Pfam; PF00481; PP2C; 1.
 DR SMART; SM00044; CYCC; 1.
 DR SMART; SM00352; PP2C; 1.
 DR SMART; SM00352; LRR_Typ; 2.
 DR SMART; SM00352; PP2C; 1.
 DR SMART; SM00352; PP2C; 1.
 DR PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
 DR PROSITE; PS50200; RA; 1.
 KW Lyase; Repeat; Leucine-rich repeat; cAMP biosynthesis; Metal-binding; Magnesium.
 FT DOMAIN 494 574 RAS-ASSOCIATING.
 FT REPEAT 632 655 LRR 1.

FT REPEAT 659 679 LRR 2.
 FT REPEAT 680 702 LRR 3.
 FT REPEAT 703 724 LRR 4.
 FT REPEAT 726 748 LRR 5.
 FT REPEAT 749 771 LRR 6.
 FT REPEAT 773 793 LRR 7.
 FT REPEAT 794 818 LRR 8.
 FT REPEAT 820 833 LRR 9.
 FT REPEAT 834 857 LRR 10.
 FT REPEAT 859 879 LRR 11.
 FT REPEAT 880 903 LRR 12.
 FT REPEAT 905 926 LRR 13.
 FT REPEAT 927 949 LRR 14.
 FT REPEAT 951 974 LRR 15.
 FT REPEAT 1004 1025 LRR 16.
 FT REPEAT 1026 1048 LRR 17.
 FT REPEAT 1050 1072 LRR 18.
 FT REPEAT 1074 1096 LRR 19.
 FT REPEAT 1101 1124 LRR 20.
 FT REPEAT 1135 1160 LRR 21.
 FT DOMAIN 1185 1440 PP2C-LIKE.
 FT DOMAIN 1441 1839 CATALYTIC (BY SIMILARITY).
 FT METAL 1488 MAGNESIUM (BY SIMILARITY).
 FT METAL 1531 MAGNESIUM (BY SIMILARITY).
 SQ SEQUENCE 1839 AA; 206895 MW; 86A69BCH1F2733CB CRC64;

Query Match 39.7%; Score 46; DB 1; Length 1839;
 Best Local Similarity 66.7%; Pred. No. 42;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 VHPVATSHUNY 14
 DB 589 VHPVATSHUNY 600

RESULT 8
 ID WR13_ARATH STANDARD; PRT; 304 AA.
 AC Q9SVB7;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Probable WRKY transcription factor 13 (WRKY DNA-binding protein 13).
 DE WRKY13 OR AY4939410 OR P23K16.40.
 GN WRKY13 OR Arabidopsis thaliana (Mouse-ear cress).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia; TISSUE=flower;
 RA Ulker B., Kushnir S., Somsich I.E.;
 RT "Arabidopsis thaliana transcription factor WRKY13.";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083488; PubMed=10617188;
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N.,
 RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
 RA Kreis M., Delseny M., Puldomenech P., Watson M., Schmidheini T.,
 RA Reichert B., Portetelle D., Perez-Alonso M., Bouty M., Bancroft I.,
 RA Vos P., Hohnsels J., Zimmermann W., Wedler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
 RA Van der Schuren J., Gyomai B., Chang Y.-J., Vandenbussche F.,
 RA Breken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Welzenegger T., Bothe G., Kampberger U., Hilbert H., Braun M.,
 RA Holzner E., Brandt A., Peters S., van Staveren J., Dikse W.,
 RA Moijman P., Klein lanthorst R., Rose M., Hauf U., Koetter P.,
 RA Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,

```

DT 16-OCT-2001 (Rel. 40, last sequence update)
DT 15-MAR-2004 (Rel. 43, last annotation update)
DE FKBP-rapamycin associated protein (FKBP).
GN FKBP OR FKBP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Bliskovsky V., Mock B.;
RT "Positional cloning of mouse plasmacytoma susceptibility gene.";
RL Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
CC -! FUNCTION: ACTS AS THE TARGET FOR THE CELL-CYCLE ARREST AND
CC IMMUNOSUPPRESSIVE EFFECTS OF THE FKBP12-RAPAMYCIN COMPLEX.
CC -! SIMILARITY: Belongs to the P13/P14-kinase family.
CC -! SIMILARITY: Contains 8 HEAT repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF152838; AAF73196.1; -.
DR HSSP; P42345; 1FAP.
DR MGD; MGI:1926394; Frap1.
DR GO; GO:0007281; P:germ-cell development; IDA.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR003151; FAT.
DR InterPro; IPR003152; FATC.
DR InterPro; IPR000357; HEAT.
DR InterPro; IPR000403; P13_P14_kinase.
DR InterPro; IPR008940; Premyl_trans.
DR pfam; PF02259; FAT; 1.
DR pfam; PF02260; FATC; 1.
DR pfam; PF00454; P13_P14_kinase; 1.
DR SMART; SMO0146; P13KC; 1.
DR PROSITE; PS00915; P13_4_KINASE_1; 1.
DR PROSITE; PS00916; P13_4_KINASE_2; 1.
DR PROSITE; PS50290; P13_4_KINASE_3; 1.
DR PROSITE; PS50077; HEAT_REPEAT; FALSE_NEG.
KM Transferase; Kinase; Repeat.
FT REPEAT 16
FT REPEAT 53 HEAT 1.
FT REPEAT 650 HEAT 2.
FT REPEAT 859 HEAT 3.
FT REPEAT 988 HEAT 4.
FT REPEAT 1069 HEAT 5.
FT REPEAT 1109 HEAT 6.
FT REPEAT 1150 HEAT 7.
FT REPEAT 1382 HEAT 8.
FT REPEAT 1933 HEAT 8.
FT DOMAIN 1282 2549 HEAT 8.
SQ SEQUENCE 2549 AA; 288734 MM; C826B36EF861BC3 CRC64;

Query Match 38.8%; Score 45; DB 1; Length 2549;
Best Local Similarity 50.0%; Pred. No. 87;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 REVHHPATSHOOQFY 16
Db 1683 RQLDHPRLTAHPQVY 1698

RESULT 10
SPMI_PIG STANDARD; PRT; 137 AA.
AC 028920;
AC 16-OCT-2001 (Rel. 40, Created)
AC 16-OCT-2001 (Rel. 40, Last sequence update)

```


"RNA polymerase II of Drosophila. Relation of its 140,000 Mr subunit to the beta subunit of Escherichia coli RNA polymerase."; RT J. Mol. Biol. 195:929-937(1987).

[2]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkley;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amaratunga P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mikos G.L.G., Abiri J.F., Agayani A., An H.-J., Andrews-Pfankuch C., Baldwin D., Bailey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Dong L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Goreil J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Laslo P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spiedling A.C., Stapleton M., Strong M., Sun B., Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weisenbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

"The genome sequence of Drosophila melanogaster."; Science 287:2185-2195(2000).

[3]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkley; TISSUE=Embryo;

RX MEDLINE=2242606; PubMed=12537569;

RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M., George R.A., Guarin H., Kronmiller B., Paclab J.M., Park S., Wan K.H., Rubin G.M., Celinker S.E.;

RA "A Drosophila full-length cDNA resource."; Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).

[4]

RP SEQUENCE OF 1-69 FROM N.A.

RC TISSUE=Embryo;

RX MEDLINE=91276237; PubMed=1905256;

RA Sitzer S., Oldenburg I., Peterson G., Bantz E.K.F.;

RT "Analysis of the promoter region of the housekeeping gene DmP140 by sequence comparison of Drosophila melanogaster and Drosophila virilis.";

RT Gene 100:155-162(1991).

CC -1- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription of DNA into RNA using the four ribonucleoside triphosphates as substrates.

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {RNA} (N).

CC -1- SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM 9 TO 14 DIFFERENT POLYPEPTIDES. RNA POLYMERASE II CONSISTS OF 10 DIFFERENT SUBUNITS. THIS SUBUNIT IS THE SECOND LARGEST COMPONENT OF RNA POLYMERASE II.

CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- MISCELLANEOUS: Three distinct zinc-containing RNA polymerases are found in eukaryotic nuclei: polymerase I for the ribosomal RNA precursor, polymerase II for the mRNA precursor, and polymerase III for 5S and tRNA genes.

CC -1- SIMILARITY: Belongs to the RNA polymerase beta chain family.

CC -1- CAUTION: Ref.1 sequence differs from that shown due to a frameshift in position 43 leading to an erroneous gene model prediction.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL, X05709; CAA29180.2; ALT FRAME.

DR EMBL, AE003703; AAF55024.1; -.

DR EMBL, EF003265; AAO25022.1; -.

DR EMBL, M62972; AAA28476.1; -.

DR PIR, A27826; A27826.

DR Flybase, FBgn003276; Rp11.40.

DR GO, GO:0005665; C:DNA-directed RNA polymerase II, core complex; NAS.

DR GO, GO:0003893; F:DNA-directed RNA polymerase activity; NAS.

DR GO, GO:0006366; P:transcription from pol II promoter; NAS.

DR InterPro, IPR007121; RNA pol B.

DR InterPro, IPR007644; RNA pol_Rpb2_1.

DR InterPro, IPR007642; RNA pol_Rpb2_2.

DR InterPro, IPR007645; RNA pol_Rpb2_3.

DR InterPro, IPR007646; RNA pol_Rpb2_4.

DR InterPro, IPR007647; RNA pol_Rpb2_5.

DR InterPro, IPR007120; RNA pol_Rpb2_6.

DR InterPro, IPR007641; RNA pol_Rpb2_7.

DR Pfam, PF04563; RNA pol_Rpb2_1; 1.

DR Pfam, PF04561; RNA pol_Rpb2_2; 1.

DR Pfam, PF04565; RNA pol_Rpb2_3; 1.

DR Pfam, PF04566; RNA pol_Rpb2_4; 1.

DR Pfam, PF04567; RNA pol_Rpb2_5; 1.

DR Pfam, PF00562; RNA pol_Rpb2_6; 1.

DR Pfam, PF04560; RNA pol_Rpb2_7; 1.

DR PROSITE, PS01166; RNA_POL_BETA; 1.

DR TRANSFERASE; DNA-directed RNA polymerase; Transcription; zinc;

DR ZINC-FINGER; Metal-binding; Nuclear protein.

FT ZN FING 1121 1142 C4-TYPE (POTENTIAL).

FT FT 72 72 A->R (IN REF.1).

FT FT 72 72

FT CONFLICT 666 667 ID->MY (IN REF.1).

FT SEQUENCE 1176 AA; 134042 MW; 224821B35BDD7F0 CRC64;

Query Match 38.4%; Score 44.5; DB 1; Length 1176;

Best Local Similarity 45.5%; Pred. No. 45;

Matches 10; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 2 EVIHP1ATSHQOQY-KIP1LV 22

DB 136 BGLDVEVTHQKFTGKIP1ML 157

RESULT 13

REF1 YEAST

ID REF1 YEAST STANDARD; PRT; 811 AA.

AC P48743;

DT 01-FEB-1996 (Rel. 33, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE RFX1-like DNA-binding protein RFX1.

GN RFX1 OR CRT1 OR YLR176C OR L9470.18.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI_TaxID=4932;

RP [1]

RP SEQUENCE FROM N.A.


```

RC STRAIN=5288c / AB972;
RX MEDLINE=97313267; PubMed=9169871.
RA Johnson M., Hillier L., Riles L., Albertman K., Andre B., Ansgorge W.,
RA Benes V., Brueckner M., Delius H., Dubois E., Duetschhoff A.,
RA Entian K.-D., Floeth M., Goffeau A., Hedling U., Hennam K.,
RA Heuss-Nelzel D., Hilbert H., Hilger F., Kleine K., Koeltter P.,
RA Louis E.-J., Messenguy F., Mews H.-W., Miesga T., Moestl D.,
RA Mueller-Auer S., Netuschil U., Obermair B., Piravandi E., Pohl T.M.,
RA Portetelle D., Purnelle B., Rechman S., Rieger M., Rinke M., Rose M.,
RA Scharte M., Scherens B., Scholler P., Schweger C., Schwarz S.,
RA Underwood A.P., Urrestarran L.A., Vandenbol M., Verhaesselt P.,
RA Vierendeels F., Voet W., Volckaert G., Voss H., Wambutt R., Weiler E.,
RA Weller H., Zimmerman F.K., Zollner A., Hani U., Hohlreis J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
CC Nature 387:87-90(1997).
CC
CC -1- SIMILARITY: Belongs to the RFX family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U17246; AAB67470.1; -.
CC DR Germonline; 142238; -.
CC DR SCD; S0004166; RFX1.
CC DR GO; GO:0005737; C:cytoplasm; IDA.
CC DR GO; GO:0005634; C:nucleus; IDA.
CC DR GO; GO:0016566; F:specific transcriptional repressor activity; IDA.
CC DR GO; GO:0000122; F:negative regulation of transcription from P. .; IDA.
CC DR InterPro; IPR003150; RFX_DNA_Binding.
CC DR Pfam; PF02257; RFX_DNA_Binding; 1.
CC KW DNA-binding.
CC
SQ SEQUENCE 811 AA; 90583 MW; 116A88B7DD84FBF0 CRC64;
-----
Oy 5 HPLATSHOQVFKTP 19
Db 496 HPLATSTYKLDPEKTP 510

Query Match 37.9%; Score 44; DB 1; Length 811;
Best Local Similarity 60.0%; Pred. No. 36;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

RESULT 14
VGLB_HSVBP STANDARD; PRT; 928 AA.
ID_VGLB_HSVBP
AC P17471:
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein I precursor (Glycoprotein GVP-6) (Glycoprotein 11A)
DE (Glycoprotein 16) (Glycoprotein G130).
GN GI.
OS Bovine herpesvirus type 1 (strain P8-2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicelloviruses.
OX NCBI_TaxID=10324;
RN 1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89020821; PubMed=2845660;
RA Mistra V., Nelson R., Smith M.;
RT "Sequence of a bovine herpesvirus type-1 glycoprotein gene that is
RT homologous to the herpes simplex gene for the glycoprotein gB.";
RL Virology 166:542-549(1988).
CC -1- SUBUNIT: Dimer, probably linked by disulfide bonds.
CC -1- SIMILARITY: Belongs to the herpesviruses glycoprotein B family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in the public domain and for commercial
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, M23257; AAA46013.1; -.
DR PIR, A31166; VGBEBG.
DR InterPro; IPRO00234; Glycoprotei_B.
DR Pfam; PF00606; Glycoprotein B; 1.
DR Prodom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 67
FT CHAIN 68 928 GLYCOPROTEIN I.
FT DOMAIN 68 780 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 781 801 POTENTIAL.
FT DOMAIN 804 824 POTENTIAL.
FT CARBOHYD 825 928 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 484 484 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 579 579 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 703 703 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 409 409 S -> T (IN REF. 1; AAA46013).
FT CONFLICT 673 673 G -> P (IN REF. 1; AAA46013).
SQ SEQUENCE 928 AA; 102177 MW; B47982224FCD769D CRC64;

Query Match 37.9%; Score 44; DB 1; Length 928;
Best Local Similarity 40.0%; Pred. No. 42;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 REVHPLATSHQRYF 15
DB 656 RELVEPCTANHKRYF 670
ID VGLB_HSVBC STANDARD; PRT; 932 AA.
VGLB_HSVBC
AC P12640;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein I precursor (Glycoprotein GVP-6) (Glycoprotein 11A)
DE (Glycoprotein 16) (Glycoprotein G130) (Glycoprotein B).
GN GI OR UL27.
OS Bovine herpesvirus type 1 (strain Cooper).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicelliovirus.
OX NCBI_TaxID=10323;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88300884; PubMed=2841484;
RA Whitbeck J.C., Bello L.J., Lawrence W.C.;
RT "Comparison of the bovine herpesvirus 1 gi gene and the herpes
RT simplex virus type 1 gB gene."
RU J. Virol. 62:3319-3327(1988).
CC -|- SUBUNIT: Dimer, probably linked by disulfide bonds.
CC -|- SIMILARITY: Belongs to the herpesviruses glycoprotein B family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, M21474; AAA46055.1; -.
DR EMBL, 278205; CAB01598.1; -.
DR EMBL, AJ004801; CA06106.1; -.
DR PIR; A28877; VGBBEC.

```

DR InterPro: IPR000234; Glycoprot_B.
 DR Pfam: PF00606; Glycoprotein_B/1.
 DR ProDom: PD000693; Glycoprot_B/1.
 KM Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 67
 FT CHAIN 1 67
 FT DOMAIN 68 932 GLYCOPROTEIN I.
 FT TRANSMEM 68 758 EXTRACELLULAR.
 FT DOMAIN 759 827 POTENTIAL.
 FT TRANSMEM 828 932 CYTOPLASMIC.
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 640 640 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 706 706 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 932 AA; 101195 MW; 9DCEA85C5FC3DA3 CRC64;

Query Match 37.9%; Score 44; DB 1; Length 932;
 Best Local Similarity 40.0%; Pred. No. 42;
 Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 REVYHPLATSHQGYF 15
 ||:|:|:|:|
 Db 659 RELVEPCTANHKXYF 673

Search completed: July 12, 2004, 14:08:02
 Job time : 6.90698 secs

RESULT 5

A;Residues: 1-675 <W00>
A;Cross-references: EMBL:AI109846: PTDN.CAB52803 1: GEPDB.CW00067. GNPB.GNP01750 01-1

C;Genetics:
A;Gene: SPDB:SPBC17G9.04c
A;Map position: 2

Query Match	40.5%;	Score 47;	DB 2;	Length 675;
Best Local Similarity	57.1%;	Pred. No. 20;		
Matches	8;	Conservative	2;	Mismatches 4;
				Indels 0;
				Gaps 0;

```
QY      3 VIHPLATSHQQYFY 16
          :| | | | | :| |
Db      566 IISPAATIHQYFFY 579
```

RESULT 8
T07623
N:Alternatn homolog HRRGP2 - soybean (fragment)
C:Species: Glycine max (soybean)
C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 11-May-2000
C:Accession: T07623
P:Hong, J.C.; Cheong, Y.H.; Nagaoka, R.T.; Bahk, J.D.; Cho, M.U.; Key, J.L.
Plant Physiol. 104, 793-796, 1994
A:Title: Isolation and characterization of three soybean extensin cDNAs.
A:Reference number: Z16058; PMID:94211912; PMID:8159793
A:Accession: T07623
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-169 <EON>
A:CROSS-references: HONB:L22030; NID:G347454; PIDN:AAA3397L.1; PID:G347455
A:Experimental source: strain Wayne; seedling
C:Genetics:
A:Gene: HRRGP2
C:Superfamily: hydroxyproline-rich glycoprotein
C:Keywords: glycoprotein; hydroxyproline

	Query Match	Similarity	Score 46;	DB 2;	length 169;
	Best Local	57.1%	Pred. No. 6;		
Matches	8;	Conservative	1;	Mismatches	5;
				Indels	0;
				Gaps	0
QY	6	PLATSHQOYFKIP	19		
	1	:			
	92	PSPISHAPPYIKSP	105		
Db					

```

transcription regulator homolog lmo0833 [imported] - listeria monocytogenes (strain EGD-
A11178
C/Species: Listeria monocytogenes
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C/Accession: A11178
R/Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fehli, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurupak, G.; Madeno, E.; Maitournam, A.; Me
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A.
Title: Comparative genomics of Listeria species.
A/Reference number: ABI077; MUID:21537279; PMID:11679669
A/Accession: A11178
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-296 <GLA>
A/Cross-references: GB:NC_003210; FIDN:CAC98911.1; FID:gl6410222; GSPDB:GN00177
A/Experimental source: strain EGD-e
C/Genetics:
A/Gene: lmo0833

```

Query	2	EVIIHPLATSHOOQFYKI	18
Query Match	39.7%	Score 46	DB 2; length 296;
Best Local Similarity	41.2%	Pred. No. 11	
Matches	7; Conservative	6; Mismatches	4; Indels 0; Gaps 0;

Db 222 EILHELSSTHDCAFYRI 238

RESULT 10
AD1149
O-acetylhomoserine sulphydrylase homolog 1m00595 [imported] - *Listeria monocytogenes* (s
C;Species: *Listeria monocytogenes*
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

Query Match	39.7%	Score 46	DB 2	Length 425
Best Local Similarity	63.6%	Pred. No. 17		
Matches 7; Conservative	3	Mismatches	1	Indels 0; Gaps 0

QY	3	VIHPLATSHQQ	13
		: : : : :	
Db	376	IHPASTTHQQ	386

RESULT 11
AD1508
O-acetylhomoserine sulfinhydrolase homolog lin0604 [imported] - *Listeria innocua* (strain
C/Species: *Listeria innocua*
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C/Accession: AD1508
R/Glasser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloekx
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussuguet, O.; Entlian, K.D.; Fshih, H
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Aubourg, Kreft, U.; Kuhn, M.; Kunst, F.; Knapkat, G.; Madheno, E.; Maltournam, A.; N
ok, C.; Schueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland
A/Title: Comparative genomics of *Listeria species*.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AD1508
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-425 <GLA>
A/Cross-references: GB:AI592022; PIRN:CA095836.1; PID:g16413044; GSPDB:GNO0178
A/Experimental source: strain C1p11262
C/genetics:
A/Gene: lin0604
C/Superfamily: O-succinylhomoserine (thiol)-lyase

	Query Match	Similarity	Score	46;	DB	2;	Length	425;
Best Local	Similarity	63.6%	Pred. No.	17;				
Matches	7;	Conservative	3;	Mismatches	1;	Indels	0;	Gaps
QY	3	VTHPLATSHQQ	13					
	:		:	:	:	:	:	:
DB	376	IITHPASTTHQQ	386					

RESULT 12
T06576
phenylalanine-tRNA ligase (EC 6.1.1.20) beta chain T32F8.180 [similarity] - Arabidopsis
C1Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 03-Jun-2002
C/Accession: T08576
R/By: M.; Zimmermann, W.; Grueneisen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; May
submitted to the Protein Sequence Database, May 1999
A/Reference number: Z16442
A/Accession: T08576
A/Molecule type: DNA
A/Residues: 1-428 <BEV>
A/Cross-references: EMBL:AL050351; GSPDB:GN00062; ATSP:T22F8.180
A/Experimental source: cultivar Columbia; BAC clone T22F8
C/Genetics:
A/Gene: ATSP:T22F8.180
A/Map position: 4
A/Intons: 49/3; 91/3; 121/3; 140/3; 179/2; 196/3; 213/1; 241/2; 273/3; 306/1; 330/1; 35
C/Superfamily: yeast cytosolic phenylalanine-tRNA ligase, beta chain
C/Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 39.7%; Score 46; DB 2; Length 428;
Best Local Similarity 46.7%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 5 HPLATSHQOYFKIP 19
|||:|:|:|:
Db 200 HPARDSDHFFELKVP 214

RESULT 13

C83975
O-acetylhomoserine sulphydrylase BH2603 [imported] - Bacillus halodurans (strain C-125)
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C/Accession: C83975
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650; WUID:20512582; PMID:11058132
A/Accession: C83975
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-430 <STO>
A/Cross-references: GB:AP001516; GB:BA000004; NID:g10175192; PIDN:BA06322.1; GSPDB:GN00
A/Experimental source: strain C-125
C/Genetics:
A/Gene: BH2603
C/Superfamily: O-succinylhomoserine (thiol)-lyase

Query Match 39.7%; Score 46; DB 2; Length 430;
Best Local Similarity 63.6%; Pred. No. 17;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 3 VHPATSHQO 13
:|:|:|:|:|:
Db 379 IHPASTHQ 389

RESULT 14

AD0341
probable membrane protein YPO2801 [imported] - Yersinia pestis (strain CO92)
C/Species: Yersinia pestis
C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C/Accession: AD0341
R/Parhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A/Reference number: AB0001; WUID:21470413; PMID:11586360
A/Accession: AD0341
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-735 <KUR>
A/Cross-references: GB:AL590842; PIDN:CA03035.1; PID:g15980773; GSPDB:GN00175
C/Genetics:

A/Gene: YPO2801

Query Match 39.7%; Score 46; DB 2; Length 735;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 3 VHPATSHQOYFKIP 22
:|:|:|:|:|:
Db 529 LHWLPDNCQOYFKIP 548

RESULT 15

C70770
hypothetical protein RV1327c - Mycobacterium tuberculosis (strain H37Rv)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C/Accession: C70770
R/Cole, S.T.; Brosch, R.; Parhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
i Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; WUID:98295987; PMID:9634230
A/Accession: C70770
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-737 <COL>
A/Cross-references: GB:E273902; GB:AL123456; NID:g3261576; PIDN:CAA98091.1; PID:e245021;
A/Experimental source: strain H37Rv
C/Genetics:
A/Gene: RV1327c

Query Match 39.7%; Score 46; DB 2; Length 737;
Best Local Similarity 46.7%; Pred. No. 32;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 5 HPLATSHQOYFKIP 19
|||:|:|:|:
Db 389 HPAREHRCQWFTLP 403

Search completed: July 12, 2004, 14:11:18
Job time : 12 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:10:33 ; Search time 38.8837 Seconds

(without alignments)
176.480 Million cell updates/sec

Title: US-10-010-667A-2_COPY_92_113
Perfect score: 116
Sequence: 1 REVHPLATSHQYFYKIPILV 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1279676 seqs, 311918243 residues

Total number of hits satisfying chosen parameters: 1279676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10C_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	116	100.0	22	14	US-10-011-095-19
2	116	100.0	22	14	US-10-010-667A-19
3	116	100.0	22	14	US-10-165-044-37
4	116	100.0	104	9	US-09-864-761-47234
5	116	100.0	267	9	US-09-747-835A-50
6	116	100.0	267	9	US-09-747-835A-51
7	116	100.0	267	12	US-10-312-312-50
8	116	100.0	267	12	US-10-312-312-51
9	116	100.0	339	9	US-09-759-143-879
10	116	100.0	339	9	US-09-780-669-879
11	116	100.0	339	9	US-09-822-827-879
12	116	100.0	339	9	US-09-802-520-11
13	116	100.0	339	9	US-09-895-793-879
14	116	100.0	339	9	US-09-895-814-879
15	116	100.0	339	12	US-10-205-267-13

16	116	100.0	339	12	US-10-408-009-2	Sequence 2, Appli
17	116	100.0	339	13	US-10-012-896-879	Sequence 879, App
18	116	100.0	339	14	US-10-011-095-2	Sequence 2, Appli
19	116	100.0	339	14	US-10-010-667A-2	Sequence 2, Appli
20	116	100.0	339	14	US-10-205-823-397	Sequence 397, App
21	116	100.0	339	14	US-10-144-678A-879	Sequence 879, App
22	116	100.0	339	14	US-10-294-025-879	Sequence 879, App
23	116	100.0	339	15	US-10-239-607-37	Sequence 37, Appl
24	116	100.0	339	15	US-10-295-027-714	Sequence 714, App
25	116	100.0	339	15	US-10-295-027-1347	Sequence 1347, App
26	116	100.0	368	12	US-10-425-114-72779	Sequence 72779, A
27	116	100.0	375	14	US-10-165-044-2	Sequence 2, Appli
28	68	58.6	26	14	US-10-165-044-40	Sequence 40, Appli
29	68	58.6	419	12	US-10-455-822-11	Sequence 11, Appli
30	68	58.6	419	12	US-10-455-822-80	Sequence 80, Appli
31	68	58.6	419	12	US-10-455-822-172	Sequence 172, App
32	68	58.6	419	12	US-10-455-822-174	Sequence 174, App
33	68	58.6	419	12	US-10-455-822-195	Sequence 195, App
34	68	58.6	419	15	US-10-239-607-32	Sequence 32, Appli
35	68	58.6	443	12	US-10-455-822-93	Sequence 93, Appli
36	68	58.6	444	12	US-10-455-822-88	Sequence 88, Appli
37	68	58.6	444	12	US-10-455-822-89	Sequence 89, Appli
38	68	58.6	444	12	US-10-455-822-90	Sequence 90, Appli
39	68	58.6	444	12	US-10-455-822-91	Sequence 91, Appli
40	68	58.6	444	12	US-10-455-822-92	Sequence 92, Appli
41	68	58.6	454	10	US-09-888-257A-10	Sequence 10, Appli
42	68	58.6	454	12	US-10-455-822-3	Sequence 3, Appli
43	68	58.6	454	12	US-10-455-822-7	Sequence 7, Appli
44	68	58.6	454	12	US-10-455-822-9	Sequence 9, Appli
45	68	58.6	454	12	US-10-455-822-19	Sequence 19, Appli

ALIGNMENTS

RESULT 1
US-10-011-095-19
; Sequence 19, Application US/10011095
; Publication No. US20030045682A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Hubert, Rene S.
; APPLICANT: Leong, Kahlan
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Saffran, Douglas C.
; APPLICANT: Mitchell, Steve Chapell
; TITLE OF INVENTION: ANTIBODIES IMMUNOSPECIFIC FOR STEAP1 (AS AMENDED)
; FILE REFERENCE: 511582001610
; CURRENT APPLICATION NUMBER: US/10/011,095
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: STEAP-1 peptide
US-10-011-095-19
Query Match 100.0%; Score 116; DB 14; Length 22;
Best Local Similarity 100.0%; Pred. No. 3.4e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 REVHPLATSHQYFYKIPILV 22
1 REVHPLATSHQYFYKIPILV 22

```
RESULT 2
US-10-010-667A-19
; Sequence 19, Application US/10010667A
; Publication No. US20030055217A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Hubert, Rene S.
; APPLICANT: Leong, Kahn
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Saffran, Douglas C.
; APPLICANT: Mitchell, Steve Chappell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 511582001601
; CURRENT APPLICATION NUMBER: US/10/010,667A
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: STEAP-1 peptide
US-10-010-667A-19

Query Match      100.0%; Score 116; DB 14; Length 22;
Best Local Similarity 100.0%; Pred. No. 3.4e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 REVHPLATSHQOYFYKIPILV 22
Db      1 REVHPLATSHQOYFYKIPILV 22

RESULT 3
US-10-165-044-37
; Sequence 37, Application US/10165044
; Publication No. US20030149531A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas Saffran
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Steven Chappell Mitchell
; APPLICANT: Mary Faris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 51158-20016.02
; CURRENT APPLICATION NUMBER: US/10/165,044
; CURRENT FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: US 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: US 60/091,183
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 09/455,486
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: WO 99/62941
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/33040
; PRIOR FILING DATE: 2000-12-06
```

```
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-165-044-37

Query Match      100.0%; Score 116; DB 14; Length 22;
Best Local Similarity 100.0%; Pred. No. 3.4e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 REVHPLATSHQOYFYKIPILV 22
Db      1 REVHPLATSHQOYFYKIPILV 22

RESULT 4
US-09-864-761-47234
; Sequence 47234, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47234
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Homo sapiens
```



```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 63
SOFTWARE: PatentIn version 3.0
SEQ ID NO 50
LENGTH: 267
TYPE: PRT
ORGANISM: Homo sapiens
US-10-312-312-50

Query Match 100.0%; Score 116; DB 12; Length 267;
Best Local Similarity 100.0%; Pred. No. 5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHOOYFKIPILV 22
Db 26 REVHPLATSHOOYFKIPILV 47

RESULT 8
US-10-312-312-51

Sequence 51, Application US/10312312
Publication No. US20040068097A1
GENERAL INFORMATION:
APPLICANT: Hysq, Inc.
APPLICANT: Yamazaki, Victoria
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Zhou, Ping
APPLICANT: Wang, Dunrui
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyun
APPLICANT: Asundi, Vinod
APPLICANT: Drmanac, Radoje T
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
FILE REFERENCE: 21272-015-061/HVS-37CIP
CURRENT FILING DATE: 2002-12-10
PRIOR FILING DATE: 2002-12-10
PRIOR APPLICATION NUMBER: US 09/729,739
PRIOR FILING DATE: 2000-12-04
PRIOR APPLICATION NUMBER: US 09/653,450
PRIOR FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: US 09/620,312
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: US 09/598,042
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: US 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 63
SOFTWARE: PatentIn version 3.0
SEQ ID NO 51
LENGTH: 267
TYPE: PRT
ORGANISM: Homo sapiens
US-10-312-312-51

Query Match 100.0%; Score 116; DB 12; Length 267;
Best Local Similarity 100.0%; Pred. No. 5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHOOYFKIPILV 22
Db 26 REVHPLATSHOOYFKIPILV 47

RESULT 9
US-09-759-143-879
Sequence 879, Application US/09759143
Patent No. US2002022248A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiaqichun

APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqi
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darriek
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C23
CURRENT FILING DATE: 2001-01-12
CURRENT APPLICATION NUMBER: US/09/759,143
CURRENT FILING DATE: 2001-01-12
NUMBER OF SEQ ID NOS: 934
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 879
LENGTH: 339
TYPE: PRT
ORGANISM: Homo sapiens
US-09-759-143-879

Query Match 100.0%; Score 116; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 6.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHOOYFKIPILV 22
Db 92 REVHPLATSHOOYFKIPILV 113

RESULT 10
US-09-780-669-879
Sequence 879, Application US/09780669
Patent No. US2002051977A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiaqichun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqi
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darriek
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William
APPLICANT: Hupel, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C24
CURRENT FILING DATE: 2001-02-09
CURRENT APPLICATION NUMBER: US/09/780,669
NUMBER OF SEQ ID NOS: 943
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 879
LENGTH: 339
TYPE: PRT

; ORGANISM: Homo sapiens
US-09-780-669-879

Query Match 100.0%; Score 116; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 6.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFYKIPILV 22
|||||
DB 92 REVHPLATSHQOYFYKIPILV 113

RESULT 11

US-09-822-827-879
; Sequence 879, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-822-827-879

Query Match 100.0%; Score 116; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 6.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFYKIPILV 22
|||||
DB 92 REVHPLATSHQOYFYKIPILV 113

RESULT 12

US-09-802-520-11
; Sequence 11, Application US/09802520
; Publication No. US20020187472A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Faris, Mary
; APPLICANT: Chen, Huel-Mei
; APPLICANT: Ison, Craig H.
; TITLE OF INVENTION: STEAP-RELATED PROTEIN
; FILE REFERENCE: PC-0037 US
; CURRENT APPLICATION NUMBER: US/09/802,520
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PERL Program
; SEQ ID NO 11
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020187472A1 g6572948
US-09-802-520-11

Query Match 100.0%; Score 116; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 6.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFYKIPILV 22
|||||
DB 92 REVHPLATSHQOYFYKIPILV 113

RESULT 13

US-09-895-793-879
; Sequence 879, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriek
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-793-879

Query Match 100.0%; Score 116; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 6.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFYKIPILV 22
|||||
DB 92 REVHPLATSHQOYFYKIPILV 113

RESULT 14

US-09-895-814-879
; Sequence 879, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriek
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.

```

; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-895-814-879

```

```

Query Match      100.0%; Score 116; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 6.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 REVHPLATSHQYFYKIPILV 22
      |||||
Db      92 REVHPLATSHQYFYKIPILV 113

```

```

RESULT 15
US-10-205-267-13
; Sequence 13, Application US/10205267
; Publication No. US20030064397A1
; GENERAL INFORMATION:
; APPLICANT: Spence, Kimberly M.
; APPLICANT: Rickett Paula K.
; APPLICANT: Lal, Preeti G.
; APPLICANT: Ison, Craig H.
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN PROSTATE AND LU
; TITLE OF INVENTION: TUMORS
; FILE REFERENCE: PV-0008 CIP
; CURRENT APPLICATION NUMBER: US/10/205,267
; CURRENT FILING DATE: 2002-07-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 339
; TYPE: PRT
; ORGANISM:
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Genbank ID No: g6572948
US-10-205-267-13

```

```

Query Match      100.0%; Score 116; DB 12; Length 339;
Best Local Similarity 100.0%; Pred. No. 6.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 REVHPLATSHQYFYKIPILV 22
      |||||
Db      92 REVHPLATSHQYFYKIPILV 113

```

Search completed: July 12, 2004, 14:27:49
 Job time : 39.8837 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:04:08 ; Search time 13.5581 Seconds

(without alignments)
83.770 Million cell updates/sec

Title: US-10-010-667a-2_COPY_92_113
Perfect score: 116
Sequence: 1 REVHPLATSHQYFYKIPILV 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	100.0	22	4	US-09-323-873A-19
2	116	100.0	339	4	US-09-323-873A-2
3	116	100.0	339	4	US-09-685-166A-879
4	46	39.7	195	4	US-09-134-001C-3548
5	44	37.9	933	2	US-08-682-847-2
6	43	37.1	337	4	US-09-252-931A-25798
7	42.5	36.6	245	4	US-09-489-039A-8972
8	42	36.2	220	4	US-09-134-000C-4003
9	42	36.2	284	3	US-09-357-251-24
10	42	36.2	319	4	US-09-134-001C-3888
11	42	36.2	1140	4	US-08-471-112A-4
12	42	36.2	1809	3	US-09-012-515A-12
13	42	36.2	1809	3	US-08-360-144A-12
14	42	36.2	1809	4	US-09-012-504A-12
15	42	36.2	1809	4	US-09-012-399A-12
16	42	36.2	2549	4	US-08-471-112A-3
17	42	36.2	2549	4	US-08-265-967C-1
18	42	36.2	2549	4	US-08-305-790B-2
19	42	36.2	2549	5	PCR-US95-06722-12
20	41.5	35.8	625	3	US-08-581-148C-18
21	41.5	35.8	625	4	US-08-759-436-3
22	41.5	35.8	625	4	US-08-759-436-5
23	41	35.3	159	4	US-09-252-991A-18982
24	41	35.3	190	1	US-08-106-981-2
25	41	35.3	322	4	US-09-540-236-2884
26	41	35.3	344	4	US-09-489-039A-8542
27	41	35.3	392	2	US-08-282-197C-60

28	41	35.3	396	3	US-09-461-474-2	Sequence 2, Appli
29	41	35.3	592	2	US-08-366-490-2	Sequence 2, Appli
30	41	35.3	592	3	US-08-860-483A-2	Sequence 2, Appli
31	41	35.3	879	1	US-08-220-151-2	Sequence 2, Appli
32	41	35.3	879	1	US-08-220-151-3	Sequence 3, Appli
33	41	35.3	879	1	US-08-413-118-2	Sequence 2, Appli
34	41	35.3	879	1	US-08-413-118-3	Sequence 3, Appli
35	41	35.3	879	1	US-08-413-118-106	Sequence 106, App
36	41	35.3	879	3	US-08-473-446-2	Sequence 2, Appli
37	41	35.3	879	3	US-08-473-446-3	Sequence 3, Appli
38	41	35.3	879	3	US-08-473-446-106	Sequence 106, App
39	40.5	34.9	261	4	US-09-107-532A-4856	Sequence 4856, Ap
40	40	34.5	68	4	US-09-134-001C-2858	Sequence 2858, Ap
41	40	34.5	77	4	US-09-328-352-7432	Sequence 7432, Ap
42	40	34.5	115	2	US-07-903-029-6	Sequence 6, Appli
43	40	34.5	206	4	US-09-219-797-2	Sequence 2, Appli
44	40	34.5	303	1	US-08-185-432-5	Sequence 5, Appli
45	40	34.5	306	4	US-09-107-532A-5464	Sequence 5464, Ap

ALIGNMENTS

```
RESULT 1
US-09-323-873A-19
; Sequence 19, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Kahan Leong
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 129,16USU2
; CURRENT APPLICATION NUMBER: US/09/323,873A
; CURRENT FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: STRAP-1 PEPTIDE
US-09-323-873A-19
Query Match 100.0%; Score 116; DB 4; Length 22;
Best Local Similarity 100.0%; Mismatch 0; Indels 0;
Matches 22; Conservative 0; Mismatches 0; Gaps 0;
QY 1 REVHPLATSHQYFYKIPILV 22
Db 1 REVHPLATSHQYFYKIPILV 22
RESULT 2
US-09-323-873A-2
; Sequence 2, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Kahan Leong
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell
```

;; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
;; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
;; FILE REFERENCE: 129 16US02
;; CURRENT APPLICATION NUMBER: US/09/323,873A
;; CURRENT FILING DATE: 1999-06-01
;; PRIOR APPLICATION NUMBER: 60/087,520
;; PRIOR FILING DATE: 1998-06-01
;; PRIOR APPLICATION NUMBER: 60/091,183
;; PRIOR FILING DATE: 1998-06-30
;; NUMBER OF SEQ ID NOS: 32
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO 2
;; LENGTH: 339
;; TYPE: PRT
;; ORGANISM: Homo Sapiens
US-09-323-873A-2

Query Match 100.0%; Score 116; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHOQYFYKIPILV 22
Db 92 REVHPLATSHOQYFYKIPILV 113

RESULT 3
US-09-685-166A-879
;; Sequence 879, Application US/09685166A
;; Patent No. 6630305
;; GENERAL INFORMATION:
;; APPLICANT: Xu, Jiangchun
;; APPLICANT: Dillon, David C.
;; APPLICANT: Mitcham, Jennifer L.
;; APPLICANT: Harlocker, Susan U.
;; APPLICANT: Uiang, Yuqun
;; APPLICANT: Henderson, Robert A.
;; APPLICANT: Kalos, Michael D.
;; APPLICANT: Fanger, Gary R.
;; APPLICANT: Retter, Marc W.
;; APPLICANT: Stolk, John A.
;; APPLICANT: Day, Craig H.
;; APPLICANT: Vedvick, Thomas S.
;; APPLICANT: Carter, Darriick
;; APPLICANT: Li, Samuel
;; APPLICANT: Wang, Aijun
;; APPLICANT: Skeiky, Yasser A.W.
;; APPLICANT: Hegler, William
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
;; FILE REFERENCE: 210121.427C21
;; CURRENT APPLICATION NUMBER: US/09/685,166A
;; CURRENT FILING DATE: 2000-10-10
;; NUMBER OF SEQ ID NOS: 898
;; SOFTWARE: FASTSEQ for Windows Version 3.0
;; SEQ ID NO 879
;; LENGTH: 339
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-685-166A-879

Query Match 100.0%; Score 116; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHOQYFYKIPILV 22
Db 92 REVHPLATSHOQYFYKIPILV 113

RESULT 4
US-09-134-001C-3548
;; Sequence 3548, Application US/09134001C

;; Patent No. 6380370
;; GENERAL INFORMATION:
;; APPLICANT: Lynn Doucette-Stamm et al
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: GTC-007
;; CURRENT APPLICATION NUMBER: US/09/134,001C
;; CURRENT FILING DATE: 1998-08-13
;; PRIOR APPLICATION NUMBER: US 60/064,964
;; PRIOR FILING DATE: 1997-11-08
;; PRIOR APPLICATION NUMBER: US 60/055,779
;; PRIOR FILING DATE: 1997-08-14
;; NUMBER OF SEQ ID NOS: 5674
;; SEQ ID NO 3548
;; LENGTH: 195
;; TYPE: PRT
;; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3548

Query Match 39.7%; Score 46; DB 4; Length 195;
Best Local Similarity 44.4%; Pred. No. 9.2;
Matches 8; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 REVHPLATSHOQYFYKI 18
Db 128 RSLSPKTSYQEMOYQI 145

RESULT 5
US-08-682-847-2
;; Sequence 2, Application US/08682847
;; Patent No. 5858989
;; GENERAL INFORMATION:
;; APPLICANT: BABLUK, LORNE
;; APPLICANT: VAN DEN HURK, SYLVIA
;; APPLICANT: ZAMB, TIM
;; APPLICANT: FITZPATRICK, DAVID
;; TITLE OF INVENTION: RECOMBINANT BOVINE HERPESVIRUS TYPE 1
;; TITLE OF INVENTION: POLYPEPTIDES AND VACCINES
;; NUMBER OF SEQUENCES: 8
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: MORRISON & ROESTER
;; STREET: 755 PAGE MILL ROAD
;; CITY: PALO ALTO
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304-1018
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/682,847
;; FILING DATE: 12-JUL-1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: PARK, FREDIE K.
;; REGISTRATION NUMBER: 35,636
;; REFERENCE/DOCKET NUMBER: 29310-20005.10
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 813-5600
;; TELEFAX: (415) 494-0792
;; TELEX: 706141
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 933 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-682-847-2

Query Match 37.9%; Score 44; DB 2; Length 933;

Best Local Similarity 40.0%; Pred. No. 1,1e+02;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 REVHPLATSHQOYF 15
||:|:|:|:|:|
Db 660 RELVEPCANMKRYF 674

RESULT 6
US-09-252-991A-25798

; Sequence 25798, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 25798

; LENGTH: 337

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-25798

Query Match 37.1%; Score 43; DB 4; Length 337;
Best Local Similarity 46.2%; Pred. No. 51;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 REVHPLATSHQO 13
||:|:|:|:|:|
Db 240 REVHPIVTHPE 252

RESULT 7
US-09-489-039A-8972

; Sequence 8972, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Bretton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; PRIOR FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 8972

; LENGTH: 245

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-8972

Query Match 36.6%; Score 42.5; DB 4; Length 245;
Best Local Similarity 35.7%; Pred. No. 43;
Matches 10; Conservative 3; Mismatches 6; Indels 9; Gaps 1;

Qy 1 REVHPLATSHQOYF 19
||:|:|:|:|:|
Db 201 RHITTEIATAKPFYVAEDDHQOYLXNP 228

RESULT 8
US-09-134-000C-4003

; Sequence 4003, Application US/09134000C

; Patent No. 6617156

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

; FILE REFERENCE: 032796-032

; CURRENT APPLICATION NUMBER: US/09/134,000C

; PRIOR FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/055,778

; PRIOR FILING DATE: 1997-08-15

; NUMBER OF SEQ ID NOS: 6812

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4003

; LENGTH: 220

; TYPE: PRT

; ORGANISM: Enterococcus faecalis

US-09-134-000C-4003

Query Match 36.2%; Score 42; DB 4; Length 220;
Best Local Similarity 41.2%; Pred. No. 45;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 3 VIHPLATSHQOYFYKIP 19
||:|:|:|:|:|
Db 153 IRYLMTKYQOEFYFEP 169

RESULT 9
US-09-357-251-24

; Sequence 24, Application US/09357251

; Patent No. 6271441

; GENERAL INFORMATION:

; APPLICANT: Falco, S. Carl

; APPLICANT: Farnodu, Layo O.

; APPLICANT: Orozco, Buddy

; APPLICANT: Schwabef, James S.

; TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetase

; FILE REFERENCE: BB-1193

; CURRENT APPLICATION NUMBER: US/09/357,251

; PRIOR FILING DATE: 1999-07-20

; EARLIER APPLICATION NUMBER: 60/093,530

; EARLIER FILING DATE: July 21, 1998

; NUMBER OF SEQ ID NOS: 37

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 24

; LENGTH: 284

; TYPE: PRT

; ORGANISM: Triticum aestivum

US-09-357-251-24

Query Match 36.2%; Score 42; DB 3; Length 284;
Best Local Similarity 46.7%; Pred. No. 60;
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 5 HPATSHQOYFYKIP 19
||:|:|:|:|:|
Db 55 HPARDSDHTEFLKAP 69

RESULT 10
US-09-134-001C-3888

; Sequence 3888, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; PRIOR FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

SEQ ID NO 3888
LENGTH: 319
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3888

Query Match 36.2%; Score 42; DB 4; Length 319;
Best Local Similarity 47.1%; Pred. No. 69;
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 4 IHPLATSHOOYFYKIP 20
DB 249 IHKXCSDHOORFYNCIS 265

RESULT 11
US-08-471-112A-4
Sequence 4, Application US/08471112A
Patent No. 6313264

GENERAL INFORMATION:
APPLICANT: Molnar-Kimber, Katherine L.
APPLICANT: Falli, Amedeo F.
APPLICANT: Caggiano, Thomas J.
APPLICANT: Nakanishi, Koji
APPLICANT: Chen, Yangu
TITLE OF INVENTION: EFFECTOR PROTEINS OF RAPAMYCIN
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner, L.L.P.
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,112A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/384,524
FILING DATE: 13-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/312,023
FILING DATE: 26-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/207,975
FILING DATE: 08-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Sleeman, Michael T.
REGISTRATION NUMBER: 36,276
REFERENCE/DOCKET NUMBER: 01142.0058-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1140 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-471-112A-4

Query Match 36.2%; Score 42; DB 4; Length 1140;
Best Local Similarity 50.0%; Pred. No. 2.8e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 REVHPLATSHOOYFY 16
DB 274 ROLDHPLPTVHPQVY 289

RESULT 12
US-09-012-515A-12
Sequence 12, Application US/09012515A
Patent No. 6127521

GENERAL INFORMATION:
APPLICANT: Berlin, Vivian
APPLICANT: Chiu, Maria Isabel
APPLICANT: Cottarel, Guillaume
APPLICANT: Damagnez, Veronique
TITLE OF INVENTION: IMMUNOSUPPRESSANT TARGET PROTEINS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/012,515A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/360,144
FILING DATE: 20-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: APV-036.02
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1809 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-012-515A-12

Query Match 36.2%; Score 42; DB 3; Length 1809;
Best Local Similarity 50.0%; Pred. No. 4.7e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 REVHPLATSHOOYFY 16
DB 943 ROLDHPLPTVHPQVY 958

RESULT 13
US-08-360-144A-12
Sequence 12, Application US/08360144A
Patent No. 6150137

GENERAL INFORMATION:
APPLICANT: Berlin, Vivian
APPLICANT: Chiu, Maria Isabel
APPLICANT: Cottarel, Guillaume
APPLICANT: Damagnez, Veronique
TITLE OF INVENTION: IMMUNOSUPPRESSANT TARGET PROTEINS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square

CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/360.144A
FILING DATE: 20-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: APV-036.02
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1809 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-360-144A-12

Query Match 36.2%; Score 42; DB 4; Length 1809;
Best Local Similarity 50.0%; Pred. No. 4.7e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 REVHPLATSHOQYFY 16
DB 943 RQDHPLEPTVHPQVY 958

RESULT 14
US-09-012-504A-12
Sequence 12, Application US/09012504A
Patent No. 6464974
GENERAL INFORMATION:
APPLICANT: Berlin, V.
APPLICANT: Chiu, I.
APPLICANT: Cottarel, G.
APPLICANT: Damagnez, V.
TITLE OF INVENTION: IMMUNOSUPPRESSANT TARGET PROTEINS
FILE REFERENCE: APRI-P05-036
CURRENT APPLICATION NUMBER: US/09/012.504A
CURRENT FILING DATE: 1998-01-23
PRIOR APPLICATION NUMBER: 08/360.144
PRIOR FILING DATE: 1994-12-20
PRIOR APPLICATION NUMBER: 08/250.795
PRIOR FILING DATE: 1994-05-27
NUMBER OF SEQ ID NOS: 35
SOFTWARE: Patentin version 3.1
SEQ ID NO 12
LENGTH: 1809
TYPE: PRT
ORGANISM: Mammalian
US-09-012-504A-12

Query Match 36.2%; Score 42; DB 4; Length 1809;
Best Local Similarity 50.0%; Pred. No. 4.7e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 REVHPLATSHOQYFY 16
DB 943 RQDHPLEPTVHPQVY 958

RESULT 15
US-09-012-399A-12

Sequence 12, Application US/09012399A
Patent No. 6509152
GENERAL INFORMATION:
APPLICANT: Berlin, Vivian
APPLICANT: Chiu, Maria Isabel
APPLICANT: Cottarel, Guillaume
APPLICANT: Damagnez, Veronique
TITLE OF INVENTION: IMMUNOSUPPRESSANT TARGET PROTEINS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/012.399A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/360.144
FILING DATE: 20-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: APV-036.02
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1809 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-012-399A-12

Query Match 36.2%; Score 42; DB 4; Length 1809;
Best Local Similarity 50.0%; Pred. No. 4.7e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 REVHPLATSHOQYFY 16
DB 943 RQDHPLEPTVHPQVY 958

Search completed: July 12, 2004, 14:12:23
Job time : 13.5581 secs

This Page Blank (uspto)

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 13:51:07 ; Search time 48.348 Seconds
(without alignments)
128.567 Million cell updates/sec

Title: US-10-010-667a-2_COPY_92_113
Perfect score: 116
Sequence: 1 REVTHPLATSHQYFYKPIIV 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Geneseqp29Jan04:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	100.0	22	3	AA58198 Human STR
2	116	100.0	22	4	AAE02786 Extracellular
3	116	100.0	104	4	AB40417 Peptide #
4	116	100.0	104	4	AAW73944 Human Don
5	116	100.0	104	4	ABG55696 Human liv
6	116	100.0	104	5	ABG43832 Human pep
7	116	100.0	254	6	ABU98426 STEAP-1 v
8	116	100.0	254	6	ABU98425 STEAP-1 v
9	116	100.0	255	6	ABU98429 STEAP-1 v
10	116	100.0	255	6	ABU98428 STEAP-1 v
11	116	100.0	258	6	ABU98395 Novel hum
12	116	100.0	258	6	ABU98389 Novel hum
13	116	100.0	258	6	ABU98424 STEAP-1 v
14	116	100.0	258	6	ABU98392 Novel hum
15	116	100.0	258	6	ABU98387 Novel hum
16	116	100.0	258	6	ABU98393 Novel hum
17	116	100.0	258	6	ABU98390 Novel hum
18	116	100.0	258	6	ABU98391 Novel hum
19	116	100.0	258	6	ABU98384 Novel hum
20	116	100.0	258	6	ABU98431 STEAP-1 v
21	116	100.0	258	6	ABU98396 Novel hum
22	116	100.0	258	6	ABU98397 Novel hum
23	116	100.0	258	6	ABU98394 Novel hum
24	116	100.0	258	6	ABU98388 Novel hum
25	116	100.0	258	6	ABU98386 Novel hum

26	116	100.0	258	6	ABU98399 Novel hum
27	116	100.0	258	6	ABU98398 Novel hum
28	116	100.0	267	6	ABU60886 Human G P
29	116	100.0	267	6	ABU60887 Human G P
30	116	100.0	282	6	ABU98432 STEAP-1 v
31	116	100.0	282	6	ABU98427 STEAP-1 v
32	116	100.0	282	6	ABU98385 Novel hum
33	116	100.0	339	3	AA58194 Human STR
34	116	100.0	339	4	AAW01282 P789P ant
35	116	100.0	339	4	AAU69927 Human pro
36	116	100.0	339	4	AAW78845 Human pro
37	116	100.0	339	4	ABU71818 Prostate
38	116	100.0	339	5	ABR95387 Human P78
39	116	100.0	339	5	ABG61813 Prostate
40	116	100.0	339	6	ABU98383 Novel hum
41	116	100.0	339	6	ABU98414 STEAP-1 v
42	116	100.0	339	6	ABU98430 STEAP-1 v
43	116	100.0	339	6	ABR54499 Prostate
44	116	100.0	339	7	ABU63313 Human six
45	116	100.0	339	7	ABD75573 Prostate

ALIGNMENTS

RESULT 1
AA58198
ID AA58198 standard; peptide; 22 AA.
XX
AC AA58198;
XX
DT 14-MAR-2000 (first entry)
XX
DE Human STRAP-1 peptide, corresponding to STRAP-1 extracellular region 1.
XX
KW Serpentine transmembrane antigen of the prostate; STRAP-1; prostate;
KW transmembrane domain; type IIIa membrane protein; expression; cancer;
KW prostate cancer; bladder cancer; colon cancer; pancreatic cancer;
KW ovarian cancer; tumour antigen; immunisation; immune response; cellular;
KW humoral; anticancer vaccine; antibody; detection; diagnosis; prognosis;
KW monitoring; susceptibility; therapeutic inhibitor; drug targeting;
KW recombinant protein.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO962941-A2.
XX
PD 09-DEC-1999.
XX
PF 01-JUN-1999; 99WO-US012157.
XX
PR 01-JUN-1998; 98US-0087520P.
XX 30-JUN-1998; 98US-0091183P.
XX
PA (UROG-) UROGENESYS INC.
PA (AFAR/) AFAR D E.
PA (HUBE/) HUBERT R S.
PA (LEON/) LEONG K.
PA (RAIT/) RAITANO A B.
PA (SAFF/) SAFFRAN D C.
XX
PI Afar DE, Hubert RS, Leong K, Raitano AB, Saffran DC;
XX WPI; 2000-072832/06.
XX
PT Novel proteins useful as diagnostic markers and therapeutic targets,
XX particularly for prostatic cancer.
XX
PS Disclosure; Page 22; 83pp; English.
XX
CC Sequences AA58198-Y58200 represent synthetic peptides that correspond to
CC the extracellular regions of STRAP-1 (serpentine transmembrane antigen of

the prostate, AAY58194). These peptides were used to raise monoclonal anti-STRAP-1 antibodies. STRAP-1 is the prototype member of the STRAP family of proteins (AAY58194-Y58197) which exhibit a high degree of structural conservation, but which show no significant structural homology to known human proteins. The STRAP-1 gene has been localised to chromosome 7p22. STRAP-1 is thought to be a type IIIa membrane protein and is expressed predominantly in prostate cells in normal human tissues. Structurally, STRAP-1 is a 339 amino acid protein characterised by six transmembrane domains and intracellular N- and C-terminal, suggesting that it folds in a "serpentine" manner into three extracellular and two intracellular loops. STRAP-1 mRNA and protein expression is maintained at high levels and throughout all stages of prostate cancer. STRAP-1 mRNA and/or protein is also overexpressed in certain other cancers, including bladder, colon, pancreatic and ovarian cancer. The function of the STRAP proteins is not known. They may be ion channels (from the presence of six transmembrane domains, a feature which is shared by certain ion channels) or gap-junction proteins (from immunohistochemical staining). STRAP-1 and STRAP-2 are cell-surface tumour antigens. Immunisation with a STRAP protein induces cellular and humoral immune responses against STRAP-expressing cells. STRAP proteins may be used to identify specific-binding agents, to produce anticancer vaccines and to generate specific antibodies. The antibodies may be used for detection, prognosis, and monitoring of cancers (or susceptibility to cancer), as therapeutic inhibitors or to target therapeutic agents to their site of action. STRAP nucleic acids may be used for recombinant protein production, as diagnostic and prognostic reagents, for identifying STRAP-expressing cells for screening inhibitors of STRAP expression and for therapeutic modulation/inhibition of STRAP expression. Since high levels of STRAP proteins are exposed on the cell surface, they are easily targeted by systemically administered agents, and because they are expressed mainly on prostatic epithelial cells, agents targeted to them should have minimal side effects on other tissues

Sequence 22 AA:

Query Match 100.0%; Score 116; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.4e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFYKPIIV 22
DB 1 REVHPLATSHQOYFYKPIIV 22

RESULT 2

AAE02786
ID AAE02786 standard; peptide; 22 AA.

AC AAE02786;

DT 06-AUG-2001 (first entry)

DE Extracellular loop #1 of human STRAP-1, suitable for cloning into pFc.

XX Human; cytosolic; antiproliferative; vaccine; gene therapy;

KW six transmembrane epithelial antigen of the prostate-1; STRAP-1;

KM chromosome 7p22.3; cancer; prostate; colon; bladder; pancreatic; lung;

KM ovarian; extracellular loop; serpentine transmembrane antigen.

XX Homo sapiens.

OS WO200140276-A2.

PN 07-JUN-2001.

PD 06-DEC-2000; 2000WO-US03040.

PR 06-DEC-1999; 99US-00455486.

XX (UROG-) UROGENESYS INC.

PA Afer DEH, Hubert RS, Raitano AB, Saffran DC, Mitchell SC;
PI Faris M, Jakobovits A;

XX WPI, 2001-367804/38.
DR
XX New STEAP (six transmembrane epithelial antigen of the prostate) PT
protein, expressed in human cancers, useful for detecting and treating PT
cancer.
XX
PS Example 19, Page 102; 187pp; English.

XX The present invention relates to human six transmembrane epithelial CC
antigen of the prostate (STEAP) protein. STEAP is a member of cell CC
surface serpentine transmembrane antigens. STEAP gene is used in gene CC
therapy. Inhibiting the development or progression of a cancer (eg. CC
prostate, colon, bladder, lung, ovarian and pancreatic) expressing STEAP CC
or inhibiting growth or killing cells expressing STEAP in a patient, CC
comprises administering a vaccine composition to the patient. Treating a CC
patient with a cancer that expresses STEAP, or inhibiting growth or CC
killing cells expressing STEAP, comprises administering to the patient a CC
vector encoding single chain monoclonal antibody that comprises the CC
variable domains of the heavy and light chains of the monoclonal antibody CC
that specifically binds to STEAP, such that the vector delivers the CC
single chain monoclonal antibody coding sequence to the cancer cells and CC
the encoded single chain monoclonal antibody is expressed CC
intracellularly. The present sequence is extracellular loop of STEAP-1 CC
suitable for cloning into pFc, which is used in the invention. STEAP-1 CC
gene is located on chromosome 7p22.3

Sequence 22 AA:

Query Match 100.0%; Score 116; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.4e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFYKPIIV 22
DB 1 REVHPLATSHQOYFYKPIIV 22

RESULT 3

ABB40417
ID ABB40417 standard; peptide; 104 AA.

AC ABB40417;

DT 04-FEB-2002 (first entry)

DE Peptide #7923 encoded by human foetal liver single exon probe.

XX Human; foetal liver; gene expression; single exon nucleic acid probe.

OS Homo sapiens.

PN WO200157277-A2.

PD 09-AUG-2001.

PR 30-JAN-2001; 2001WO-US000669.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

Human genome-derived single exon nucleic acid probes useful for analyzing

PT gene expression in human fetal liver.
XX
PS Claim 27; SEQ ID NO 33052; 639pp + Sequence Listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 104 AA;

Query Match 100.0%; Score 116; DB 4; Length 104;
Best Local Similarity 100.0%; Pred. No. 7.4e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFYKIPILV 22
Db 59 REVHPLATSHQOYFYKIPILV 80

RESULT 4
AAM73944
ID AAM73944 standard; protein; 104 AA.
XX
AC AAM73944;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 34250.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
FN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-00608408.
XX
PR 03-AUG-2000; 2000US-00632366.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 34250; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention
XX
SQ Sequence 104 AA;

Query Match 100.0%; Score 116; DB 4; Length 104;
Best Local Similarity 100.0%; Pred. No. 7.4e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFYKIPILV 22
Db 59 REVHPLATSHQOYFYKIPILV 80

RESULT 5
ABG55696
ID ABG55696 standard; peptide; 104 AA.
XX
AC ABG55696;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver peptide, SEQ ID NO 34344.
XX
KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
KW hypercholesterolaemia; coronary heart disease.
XX
OS Homo sapiens.
XX
FN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000664.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-00608408.
XX
PR 03-AUG-2000; 2000US-00632366.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human adult liver.
XX
PS Claim 27; SEQ ID NO 34344; 658pp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (II) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABG47348-ABG59930 represent human
CC liver single exon encoded peptides of the invention. Note: The sequence
CC information for this patent does not appear in the printed specification
CC but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 104 AA;

Query Match 100.0%; Score 116; DB 4; Length 104;
Best Local Similarity 100.0%; Pred. No. 7.4e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQOYFYKIPILV 22
Db 59 REVHPLATSHQOYFYKIPILV 80

RESULT 6
ABG43832
ID ABG43832 standard; peptide; 104 AA.
XX AC ASG43832;
XX XX 19-AUG-2002 (first entry)
DE Human peptide encoded by genome-derived single exon probe SEQ ID 33497.
XX XX
KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberos sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hemansky-Pudlak syndrome; sarcoidosis; pulmonary haemostasis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagazer syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KM hyaline membrane disease.
XX OS Homo sapiens.
XX XX
ED W0200186003-A2.
XX XX
PD 15-NOV-2001.
XX PF
PR 30-JAN-2001; 2001WO-US00065.
XX PR
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX XX
EA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WP; 2002-114183/15.
PT Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples.
PS Claim 27; SEQ ID NO 33497; 634BP; English.
XX XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC / the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray, assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the

CC epithelial antigen of the prostate or STEAP-1
XX
SQ Sequence 254 AA;

Query Match 100.0%; Score 116; DB 6; Length 254;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQGYFYKIPILV 22
Db 92 REVHPLATSHQGYFYKIPILV 113

RESULT 8
ABU98425
ID ABU98425 standard; protein; 254 AA.

XX ABU98425;

XX 31-JUL-2003 (first entry)

DE STEAP-1 variant 8P1D4 v.1 #1.

XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX vaccine.

OS Homo sapiens.

PN W02003022995-A2.

PD 20-MAR-2003.

PF 06-SEP-2002; 2002WO-US028371.

PR 06-SEP-2001; 2001US-0317840P.

PR 05-APR-2002; 2002US-0370387P.

XX (AGEN-) AGENSYS INC.

PI Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;

DR WPI; 2003-313240/30.

PT New composition comprising a substance that modulates the status of a
STEAP-1-related protein, useful for treating and detecting cancer.

XX Example 53; Page 169-170; 248pp; English.

CC The invention describes a composition comprising a substance that
modulates the status of a protein (I) of 340 or 283 amino acids, or of
any of the 15 sequences of 259 amino acids, given in the specification,
or a molecule that is modulated by the protein, where the status of the
cell that expresses the protein is modulated. The compositions, proteins,
polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
precision, cytogenetic abnormalities in the chromosomal region that
encodes STEAP-1 that may contribute to the malignant phenotype. This is
the amino acid sequence of a variant of human six transmembrane
epithelial antigen of the prostate or STEAP-1

XX Sequence 254 AA;

Query Match 100.0%; Score 116; DB 6; Length 254;

Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQGYFYKIPILV 22
Db 92 REVHPLATSHQGYFYKIPILV 113

RESULT 9
ABU98429
ID ABU98429 standard; protein; 255 AA.

XX ABU98429;

XX 31-JUL-2003 (first entry)

DE STEAP-1 variant 8P1D4 v.3 #2.

XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX vaccine.

OS Homo sapiens.

PN W02003022995-A2.

PD 20-MAR-2003.

PF 06-SEP-2002; 2002WO-US028371.

PR 06-SEP-2001; 2001US-0317840P.

PR 05-APR-2002; 2002US-0370387P.

XX (AGEN-) AGENSYS INC.

PI Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;

DR WPI; 2003-313240/30.

PT New composition comprising a substance that modulates the status of a
STEAP-1-related protein, useful for treating and detecting cancer.

XX Example 53; Page 172; 248pp; English.

CC The invention describes a composition comprising a substance that
modulates the status of a protein (I) of 340 or 283 amino acids, or of
any of the 15 sequences of 259 amino acids, given in the specification,
or a molecule that is modulated by the protein, where the status of the
cell that expresses the protein is modulated. The compositions, proteins,
polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
precision, cytogenetic abnormalities in the chromosomal region that
encodes STEAP-1 that may contribute to the malignant phenotype. This is
the amino acid sequence of a variant of human six transmembrane
epithelial antigen of the prostate or STEAP-1

XX Sequence 255 AA;

Query Match 100.0%; Score 116; DB 6; Length 255;

Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQGYFYKIPILV 22
Db 92 REVHPLATSHQGYFYKIPILV 113

RESULT 10

ABU98428
ID ABU98428 standard; protein; 255 AA.

XX ABU98428;

XX 31-JUL-2003 (first entry)

DE STEAP-1 variant 8P1D4 v.2 #3.

XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX vaccine.

```
XX OS Homo sapiens.
XX PN WO2003022995-A2.
XX PD 20-MAR-2003.
XX PF 06-SEP-2002; 2002WO-US028371.
XX PR 06-SEP-2001; 2001US-0317840P.
XX PR 05-APR-2002; 2002US-0370387P.
XX PA (AGEN-) AGENSYS INC.
XX PI Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
XX WPI; 2003-313240/30.
XX DR WPI; 2003-313240/30.
XX PT New composition comprising a substance that modulates the status of a
XX PT STEAP-1-related protein, useful for treating and detecting cancer.
XX PS Example 53; Page 172; 248pp; English.
XX CC The invention describes a composition comprising a substance that
XX CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
XX CC any of the 15 sequences of 259 amino acids, given in the specification,
XX CC or a molecule that is modulated by the protein, where the status of the
XX CC cell that expresses the protein is modulated. The compositions, proteins,
XX CC polynucleotides and methods are useful for treating and detecting cancer.
XX CC The STEAP-1-related proteins are useful for generating cancer vaccines.
XX CC The polynucleotides are useful as tools for delineating, with greater
XX CC precision, cytogenetic abnormalities in the chromosomal region that
XX CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
XX CC the amino acid sequence of a variant of human six transmembrane
XX CC epithelial antigen of the prostate or STEAP-1
XX SQ Sequence 255 AA;

Query Match 100.0%; Score 116; DB 6; Length 255;
Best Local Similarity 100.0%; Pred. No. 2e-10; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

QY 1 REVHPLATSHQOYFYKIPILV 22
Db 92 REVHPLATSHQOYFYKIPILV 113

RESULT 11
ABU98395 ID ABU98395 standard; protein; 258 AA.
XX AC ABU98395;
XX DT 31-JUL-2003 (first entry)
XX DE Novel human gene STEAP-1 variant 13.
XX KM STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
XX KM cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX KM vaccine.
XX OS Homo sapiens.
XX PN WO2003022995-A2.
XX PD 20-MAR-2003.
XX PF 06-SEP-2002; 2002WO-US028371.
XX PR 06-SEP-2001; 2001US-0317840P.
XX PR 05-APR-2002; 2002US-0370387P.
XX PA (AGEN-) AGENSYS INC.
```

```
XX PI Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
XX WPI; 2003-313240/30.
XX DR WPI; 2003-313240/30.
XX DR N-PSDB; ACD02603.
XX PT New composition comprising a substance that modulates the status of a
XX PT STEAP-1-related protein, useful for treating and detecting cancer.
XX PS Example 2; Fig 2K; 248pp; English.
XX CC The invention describes a composition comprising a substance that
XX CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
XX CC any of the 15 sequences of 259 amino acids, given in the specification,
XX CC or a molecule that is modulated by the protein, where the status of the
XX CC cell that expresses the protein is modulated. The compositions, proteins,
XX CC polynucleotides and methods are useful for treating and detecting cancer.
XX CC The STEAP-1-related proteins are useful for generating cancer vaccines.
XX CC The polynucleotides are useful as tools for delineating, with greater
XX CC precision, cytogenetic abnormalities in the chromosomal region that
XX CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
XX CC the amino acid sequence of a variant of human six transmembrane
XX CC epithelial antigen of the prostate or STEAP-1
XX SQ Sequence 258 AA;

Query Match 100.0%; Score 116; DB 6; Length 258;
Best Local Similarity 100.0%; Pred. No. 2e-10; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

QY 1 REVHPLATSHQOYFYKIPILV 22
Db 92 REVHPLATSHQOYFYKIPILV 113

RESULT 12
ABU98389 ID ABU98389 standard; protein; 258 AA.
XX AC ABU98389;
XX DT 31-JUL-2003 (first entry)
XX DE Novel human gene STEAP-1 variant 7.
XX KM STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
XX KM cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX KM vaccine.
XX OS Homo sapiens.
XX PN WO2003022995-A2.
XX PD 20-MAR-2003.
XX PF 06-SEP-2002; 2002WO-US028371.
XX PR 06-SEP-2001; 2001US-0317840P.
XX PR 05-APR-2002; 2002US-0370387P.
XX PA (AGEN-) AGENSYS INC.
XX PI Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
XX WPI; 2003-313240/30.
XX DR N-PSDB; ACD02603.
XX PT New composition comprising a substance that modulates the status of a
XX PT STEAP-1-related protein, useful for treating and detecting cancer.
XX PS Example 2; Fig 2G; 248pp; English.
XX CC The invention describes a composition comprising a substance that
```


modulates the status of a protein (1) of 340 or 283 amino acids, or of any of the 15 sequences of 259 amino acids, given in the specification, or a molecule that is modulated by the protein, where the status of the cell that expresses the protein is modulated. The compositions, proteins, polynucleotides and methods are useful for treating and detecting cancer. The STEAP-1-related proteins are useful for generating cancer vaccines. The polynucleotides are useful as tools for delineating, with greater precision, cytogenetic abnormalities in the chromosomal region that encodes STEAP-1 that may contribute to the malignant phenotype. This is the amino acid sequence of a variant of human six transmembrane epithelial antigen of the prostate or STEAP-1

SQ Sequence 258 AA;

Query Match 100.0%; Score 116; DB 6; Length 258;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQGYFYKIPILV 22
|||
Db 92 REVHPLATSHQGYFYKIPILV 113

RESULT 13

ABU98424 ABU98424 standard; protein; 258 AA.

XX ABU98424;

AC 31-JUL-2003 (first entry)

DT STEAP-1 variant 8P1D4 V.2 #1.

DE STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;

KW cancer vaccine; delineation; cytogenetic abnormality; cyrostatic;

KM vaccine.

XX Homo sapiens.

OS WO2003022995-A2.

PN 20-MAR-2003.

PD 06-SEP-2002; 2002WO-US028371.

PF 06-SEP-2001; 2001US-0317840P.

PR 05-APR-2002; 2002US-0370387P.

PS (AGEN-) AGENSYS INC.

PI Faris M, Ge W, Raitano AB, Chalilita-Eid PM, Jakobovits A;

XX WPI; 2003-313240/30.

DR New composition comprising a substance that modulates the status of a

PT STEAP-1-related protein, useful for treating and detecting cancer.

XX Example 53; Page 169; 248pp; English.

PS The invention describes a composition comprising a substance that

CC modulates the status of a protein (1) of 340 or 283 amino acids, or of

CC any of the 15 sequences of 259 amino acids, given in the specification,

CC or a molecule that is modulated by the protein, where the status of the

CC cell that expresses the protein is modulated. The compositions, proteins,

CC polynucleotides and methods are useful for generating cancer vaccines.

CC The STEAP-1-related proteins are useful as tools for delineating, with greater

CC precision, cytogenetic abnormalities in the chromosomal region that

CC encodes STEAP-1 that may contribute to the malignant phenotype. This is

CC the amino acid sequence of a variant of human six transmembrane

CC epithelial antigen of the prostate or STEAP-1

CC Sequence 258 AA;

SQ

Query Match 100.0%; Score 116; DB 6; Length 258;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQGYFYKIPILV 22

Db 92 REVHPLATSHQGYFYKIPILV 113

RESULT 14

ABU98392 ABU98392 standard; protein; 258 AA.

XX ABU98392;

AC 31-JUL-2003 (first entry)

DT Novel human gene STEAP-1 variant 10.

DE STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;

KW cancer vaccine; delineation; cytogenetic abnormality; cyrostatic;

KM vaccine.

XX Homo sapiens.

OS WO2003022995-A2.

PN 20-MAR-2003.

PD 06-SEP-2002; 2002WO-US028371.

PF 06-SEP-2001; 2001US-0317840P.

PR 05-APR-2002; 2002US-0370387P.

PS (AGEN-) AGENSYS INC.

PI Faris M, Ge W, Raitano AB, Chalilita-Eid PM, Jakobovits A;

XX WPI; 2003-313240/30.

DR N-PSDB; ACDD02606.

PD New composition comprising a substance that modulates the status of a

PT STEAP-1-related protein, useful for treating and detecting cancer.

XX Example 2; Fig 2J; 248pp; English.

PS The invention describes a composition comprising a substance that

CC modulates the status of a protein (1) of 340 or 283 amino acids, or of

CC any of the 15 sequences of 259 amino acids, given in the specification,

CC or a molecule that is modulated by the protein, where the status of the

CC cell that expresses the protein is modulated. The compositions, proteins,

CC polynucleotides and methods are useful for generating cancer vaccines.

CC The STEAP-1-related proteins are useful as tools for delineating, with greater

CC precision, cytogenetic abnormalities in the chromosomal region that

CC encodes STEAP-1 that may contribute to the malignant phenotype. This is

CC the amino acid sequence of a variant of human six transmembrane

CC epithelial antigen of the prostate or STEAP-1

CC Sequence 258 AA;

SQ

Query Match 100.0%; Score 116; DB 6; Length 258;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQGYFYKIPILV 22

Db 92 REVHPLATSHQGYFYKIPILV 113

RESULT 15

ABU98387

ID ABU98387 standard; protein; 258 AA.
XX
AC ABU98387;
XX
DT 31-JUL-2003 (first entry)
XX
DE Novel human gene STEAP-1 variant 5.
XX
KW STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
KW cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
KW vaccine.
XX
OS Homo sapiens.
XX
PN WO2003022995-A2.
XX
PD 20-MAR-2003.
XX
PF 06-SEP-2002; 2002WO-US028371.
XX
PR 06-SEP-2001; 2001US-0317840P.
PR 05-APR-2002; 2002US-0370387P.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Faris M, Ge W, Raitano AB, Challita-Eid PM, Jakobovits A;
XX
DR WPI; 2003-313240/30.
DR N-PSDB; ACD02601.
XX
PT New composition comprising a substance that modulates the status of a
PT STEAP-1-related protein, useful for treating and detecting cancer.
XX
PS Example 2; Fig 2E; 248pp; English.
XX
CC The invention describes a composition comprising a substance that
CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
CC any of the 15 sequences of 259 amino acids, given in the specification,
CC or a molecule that is modulated by the protein, where the status of the
CC cell that expresses the protein is modulated. The compositions, proteins,
CC polynucleotides and methods are useful for generating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
XX
SQ Sequence 258 AA;
XX
Query Match 100.0%; Score 116; DB 6; Length 258;
Best Local Similarity 100.0%; Pred. No. 2e-10; Mismatches 0; Gaps 0;
Matches 22; Conservative 0; Indels 0;
OY 1 REVHPLATSHQOYFYKIPILV 22
DB 92 REVHPLATSHQOYFYKIPILV 113
Search completed: July 12, 2004, 14:07:22
Job time : 48.3488 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:00:58 ; Search time 22.1512 Seconds

(Without alignments)
213.658 Million cell updates/sec

Title: US-10-010-667a-2_COPY_277_291

Perfect score: 93

Sequence: 1 WIDIKQFWYTPPTF 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	97.8	338	6	Q9GL50
2	80	86.0	339	11	Q924J9
3	80	86.0	339	11	Q924J2
4	80	86.0	339	11	Q9CWR7
5	55	59.1	419	16	Q8F0P8
6	49	52.7	448	16	Q98CK3
7	48	51.6	246	15	Q8UHF2
8	48	51.6	453	16	Q8UHF8
9	46	49.5	122	15	Q9IUL9
10	46	49.5	844	15	Q97002
11	45	48.4	107	2	Q8EYQ0
12	45	48.4	107	9	Q8FAJ1
13	45	48.4	122	15	Q90DN6
14	45	48.4	122	15	Q9QIY7
15	45	48.4	122	15	Q9YXR4
16	45	48.4	122	15	Q9YXQ1

17	45	48.4	122	15	Q9IIL6	Q9IIL6 human immun
18	45	48.4	122	15	Q9QIU8	Q9QIU8 human immun
19	45	48.4	122	15	Q7ZUR8	Q7ZUR8 human immun
20	45	48.4	122	15	Q7ZUR5	Q7ZUR5 human immun
21	45	48.4	133	15	Q8UQZ0	Q8UQZ0 human immun
22	45	48.4	133	15	Q8UQY9	Q8UQY9 human immun
23	45	48.4	133	15	Q8UR20	Q8UR20 human immun
24	45	48.4	151	15	Q7SM24	Q7SM24 human immun
25	45	48.4	153	15	Q7SM42	Q7SM42 human immun
26	45	48.4	155	15	Q8UJQ3	Q8UJQ3 human immun
27	45	48.4	161	15	Q8JER1	Q8JER1 human immun
28	45	48.4	163	15	Q90E72	Q90E72 human immun
29	45	48.4	167	15	Q8JER2	Q8JER2 human immun
30	45	48.4	169	15	Q8JER0	Q8JER0 human immun
31	45	48.4	169	15	Q8JER9	Q8JER9 human immun
32	45	48.4	214	15	Q8AU20	Q8AU20 human immun
33	45	48.4	358	15	Q78120	Q78120 human immun
34	45	48.4	375	11	Q8BUH3	Q8BUH3 mus musculus
35	45	48.4	483	16	Q82Z89	Q82Z89 enterococcus
36	45	48.4	711	15	Q7SVK4	Q7SVK4 human immun
37	45	48.4	711	15	Q7SVK3	Q7SVK3 human immun
38	45	48.4	713	15	Q7SVK2	Q7SVK2 human immun
39	45	48.4	790	15	Q8Q2X3	Q8Q2X3 human immun
40	45	48.4	833	15	Q9QKI4	Q9QKI4 human immun
41	45	48.4	833	15	Q9QKH9	Q9QKH9 human immun
42	45	48.4	833	15	Q9QKI3	Q9QKI3 human immun
43	45	48.4	840	15	Q8Q2X4	Q8Q2X4 human immun
44	45	48.4	841	15	Q9IV31	Q9IV31 human immun
45	45	48.4	846	15	Q89292	Q89292 human immun

ALIGNMENTS

RESULT 1

ID Q9GL50 PRELIMINARY: PRT: 338 AA.

AC Q9GL50: PRELIMINARY: PRT: 338 AA.
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Six transmembrane endothelial antigen of PANC.
 OS Sus scrofa (pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 CC NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nagasaka T., Boulday G., Coupel S., Coulon F., Tesson L.,
 RA Heslan J.-M., Soullion J.-P., Chaireau B.,
 RT "Differential gene expression in endothelial cells during TNF-alpha-
 RT and lip-mediated activation."
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF319659; AAC33868.1; -;
 DR GO: GO:0016021; C: integral to membrane; IEA.
 KW Transmembrane.
 SQ SEQUENCE 338 AA: 39918 MW; ED4908E067A32B CRC64;

Query Match 97.8%; Score 91; DB 6; Length 338;
 Best Local Similarity 86.7%; Pred. No. 1,1e-06;
 Matches 13; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WIDIKQFWYTPPTF 15
 DB 276 WVDIKQFWYTPPTF 290
 RESULT 2
 ID Q924J9 PRELIMINARY: PRT: 339 AA.
 AC Q924J9:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

```
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Six-transmembrane epithelial antigen of the prostate.
GN STEAP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=21371909; PubMed=11479226;
RA Yang D., Holt G.E., Velchers M.P., Kwon E.D., Kast W.M.;
RT "Murine six-transmembrane epithelial antigen of the prostate, prostate
RT stem cell antigen, and prostate-specific membrane antigen: prostate-
RT specific cell-surface antigens highly expressed in prostate cancer of
RT transgenic adenocarcinoma mouse prostate mice.";
RL Cancer Res. 61:5857-5860(2001).
DR EMBL; AF297098; AAK83126.1; -.
DR MGD; MGI:1917608; Steap.
DR GO; GO:0016021; C:integral to membrane; IEA.
KW Transmembrane.
SQ SEQUENCE 339 AA; 39318 MW; 4B26A71FF559E84F CRC64;

Query Match 86.0%; Score 80; DB 11; Length 339;
Best Local Similarity 73.3%; Pred. No. 6.6e-05;
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WIDIKQFWYTPPTF 15
Db 277 WVDVQFWYMPPTF 291

RESULT 3
ID 0924Z2 PRELIMINARY; PRT; 339 AA.
AC 0924Z2;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Dudaliln.
GN STEAP OR 1010001D01RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Seru V., Manivet P., Lambdin D., Vaubourdoille M., Kellermann O.,
RA Loric S.;
RT "Prostate and non-prostate expression of dudaliln, the mouse ortholog
RT of human STEAP.";
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY029584; AAK50537.1; -.
DR MGD; MGI:1917608; Steap.
SQ SEQUENCE 339 AA; 39109 MW; 32A2C29F2E333BD0 CRC64;

Query Match 86.0%; Score 80; DB 11; Length 339;
Best Local Similarity 73.3%; Pred. No. 6.6e-05;
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WIDIKQFWYTPPTF 15
Db 277 WVDVQFWYMPPTF 291

RESULT 4
ID 09CWR7 PRELIMINARY; PRT; 339 AA.
AC 09CWR7;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE 2410007B19RIK protein.
```

```
GN STEAP OR 2410007B19RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai U., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa T., Hara A., Fukunishi Y., Konno H., Adachi U., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staib F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake U., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya K., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyokaka K., Wang K.H., Weitz C., Wittkeker C., Wilming L.,
RA Wyszew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010437; BAB26938.1; -.
DR MGD; MGI:1917608; Steap.
SQ SEQUENCE 339 AA; 39264 MW; 3F7AB9C7520F0968 CRC64;

Query Match 86.0%; Score 80; DB 11; Length 339;
Best Local Similarity 73.3%; Pred. No. 6.6e-05;
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WIDIKQFWYTPPTF 15
Db 277 WVDVQFWYMPPTF 291

RESULT 5
ID 08F0P8 PRELIMINARY; PRT; 419 AA.
AC 08F0P8;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Conserved hypothetical protein.
GN IA3444.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=1173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011502; AAN50642.1; -.
DR InterPro; IPR001466; Beta_lactamase.
DR InterPro; IPR000437; Prok_1lipoprot_S.
DR Pfam; PF00144; beta-lactamase; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 419 AA; 47652 MW; BE6FE7CF180F0B80 CRC64;

Query Match 59.1%; Score 55; DB 16; Length 419;
Best Local Similarity 50.0%; Pred. No. 0.99;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 WIDIKQFWYTPPT 14
Db 277 WVDVQFWYMPPT 291
```

Db 248 WIDFKDYEWYSPS 261

RESULT 6

Q98CK3 PRELIMINARY; PRT; 448 AA.

AC Q98CK3; 01-OCT-2001 (TREMBlrel. 18, Created)

DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)

DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

GN Alpha-glucosidase ABC transporter, substrate-binding protein. ML5113.

OS Rhizobium loti (Mesorhizobium loti).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Mesorhizobium.

NC NCBITaxID=381;

Sequence from N.A.

STRAIN=MAF030399;

Medline=21082930; Pubmed=11214968;

Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S., Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T., Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;

"Complete genome structure of the nitrogen-fixing symbiotic bacterium Mesorhizobium loti."

DNA Res. 7:331-338(2000).

EMBL; AP003005; BAB51618.1;

DR GO; GO:0005215; P:transporter activity; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR006059; SBP_bac_1.

PFam; PF01547; SBP_bac_1; 1.

Complete proteome.

Sequence 448 AA; 48920 MW; 987B8569A945D19 CRC64;

Query Match 52.7%; Score 49; DB 16; Length 448;

Best Local Similarity 53.8%; Pred. No. 10;

Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 DIKQFWYTPPTF 15

Db 164 DVKSLWYSPDNF 176

RESULT 7

Q8JF92 PRELIMINARY; PRT; 246 AA.

AC Q8JF92; 01-OCT-2002 (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE Truncated envelope protein (Fragment).

OS Human immunodeficiency virus 1.

OC Viruses; Retroviridae; Lentivirus.

NC NCBITaxID=11676;

Sequence from N.A.

STRAIN=RF;

Medline=22011541; Pubmed=12015905;

DR Jones D.R., Suzuki K., Piller S.C.;

"A 100-Amino Acid Truncation in the Cytoplasmic Tail of Glycoprotein 41 in the Reference HIV Type 1 Strain RF."

ALMS Res. Hum. Retroviruses 18:513-517(2002).

EMBL; AF403705; AAM21476.1;

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0019031; C:lipid envelope; IEA.

DR GO; GO:0005198; F:structural molecule activity; IEA.

InterPro; IPR000328; Env_GP41.

PFam; PF00517; GP41; 1.

Envelope protein; Transmembrane.

NON_TER 1

Sequence 246 AA; 27734 MW; 4F0058F96DE1C4B3 CRC64;

Query Match 51.6%; Score 48; DB 15; Length 246;

Best Local Similarity 70.0%; Pred. No. 8;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WIDIKQFWY 10

Db 162 WEDITQWY 171

RESULT 8

Q8JHT8 PRELIMINARY; PRT; 453 AA.

AC Q8JHT8; 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DE ABC transporter, substrate binding protein.

GN AGLE OR AFU0591 OR AGR_C1045.

OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.

NC NCBITaxID=176299;

Sequence from N.A.

Medline=21608550; Pubmed=11743193;

Wood D.W., Sebubal J.C., Kaul R., Monks D.E., Kitajima J.P., Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L., Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr., Chapman P., Clendinning J., Deatherage G., Gillet M., Grant C., Kuyavlin T., Levy R., Li M.-J., McClelland E., Palmieri A., Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D., Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan M., Perry M., Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M., Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V., Nester E.W.;

"The genome of the natural genetic engineer Agrobacterium tumefaciens C58."

Science 294:2317-2323(2001).

Sequence from N.A.

Medline=21608551; Pubmed=11743194;

Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Qurollo B., Goldman B.S., Cao Y., Askenzi M., Halling C., Mullin L., Houtmel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F., Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G., Cielo C., Slater S.;

"Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58."

Science 294:2323-2328(2001).

EMBL; AE009027; AAL1608.1;

DR EMBL; AE007993; AAK86402.1;

DR PIR; A97431; A97431.

DR PIR; AB2649; AB2649.

DR GO; GO:0005215; P:transporter activity; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR006059; SBP_bac_1.

PFam; PF01547; SBP_bac_1; 1.

Complete proteome.

Sequence 453 AA; 49090 MW; 90896249313CD85E CRC64;

Query Match 51.6%; Score 48; DB 16; Length 453;

Best Local Similarity 53.8%; Pred. No. 15;

Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 DIKQFWYTPPTF 15

Db 169 DVKSLWYVPENF 181

RESULT 9

Q9IUI9

ID Q91TL9 PRELIMINARY; PRT; 122 AA.
AC Q91TL9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AR3;
RX MEDLINE=20346416; PubMed=10890362;
RA Masciotra S., Liellera B., Belloso W., Clara L., Tanuri A., Ramos A.,
Baggi J., Lal R., Pieniazek D.;
RT "Evidence for a high frequency of HIV-1 subtype F infections among
RT heterosexual population in Buenos Aires, Argentina.";
RL AIDS Res. Hum. Retroviruses 16:1007-1014(2000).
DR EMBL: AF220713; AAF6832.1; -;
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
DR InterPro: IPR000328; F:structural molecule activity; IEA.
DR Pfam: PF00517; GP41; 1. _GP41.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 122
SQ SEQUENCE 122 AA; 14840 MW; 10821A3BCC39F4E6 CRC64;

Query Match 49.5%; Score 46; DB 15; Length 122;
Best Local Similarity 60.0%; Pred. No. 8.3;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WIDIKQFWY 10
Db 112 WEDISQWLMY 121

RESULT 10
Q97002 PRELIMINARY; PRT; 844 AA.
AC Q97002;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HIVBR020.17;
RX MEDLINE=96190564; PubMed=8627686;
RA Gao F., Morrison S.G., Robertson D.L., Thornton C.L., Craig S.,
Karlsson G., Sodroski J., Morgado M., Galvao-Castro B.,
von Briesen H., Beddows S., Weber J., Sharp P.M., Shaw G.M.,
Hahn B.H.;
RT "Molecular cloning and analysis of functional envelope genes from
RT human immunodeficiency virus type 1 sequence subtypes A through G. The
RT WHO and NIAID Networks for HIV Isolation and Characterization.";
RL J. Virol. 70:1651-1657(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HIVBR020.17;
RA Allan E.E.;
RT Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: U27401; AAB06242.1; -;
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.

DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 844 AA; 95594 MW; 0B02E379CEFD17691 CRC64;

Query Match 49.5%; Score 46; DB 15; Length 844;
Best Local Similarity 60.0%; Pred. No. 59;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WIDIKQFWY 10
Db 660 WEDISQWLMY 669

RESULT 11
Q9EVO0 PRELIMINARY; PRT; 107 AA.
AC Q9EVO0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E32511;
RX MEDLINE=20407286; PubMed=10948097;
RA Unkneif A., Schmidt H.;
RT "Structural analysis of phage-borne stx genes and their flanking
RT sequences in shiga toxin-producing escherichia coli and shiga11a
RT dysenteriae type 1 strains.";
RL Infect. Immun. 68:4856-4864(2000).
DR EMBL: AJ251452; CAC05542.1; -;
KW Hypothetical protein.
SQ SEQUENCE 107 AA; 12177 MW; 8BCD6C52B7D29CB6 CRC64;

Query Match 48.4%; Score 45; DB 2; Length 107;
Best Local Similarity 46.7%; Pred. No. 11;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WIDIKQFWYTPPTF 15
Db 22 WCNIKILIMYQSDTF 36

RESULT 12
Q8HAJ1 PRELIMINARY; PRT; 107 AA.
AC Q8HAJ1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Bacteriophage LC159.
OC Viruses.
OX NCBI_TaxID=210928;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=159;
RA Muniesa M., Jofre J.;
RT "Variability of shiga converting bacteriophages in E. coli O157:H7
RT strains of human origin isolated from the same outbreak.";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF548456; AAN59922.1; -;
KW Hypothetical protein.
SQ SEQUENCE 107 AA; 12197 MW; 8BCD6C52B7D29CB6 CRC64;

Query Match 48.4%; Score 45; DB 9; Length 107;
Best Local Similarity 46.7%; Pred. No. 11;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WIDIKOFVWYTPPTF 15
Db 22 WCNKILIMYQSDTF 36

RESULT 13

Q90DN6 PRELIMINARY; PRT; 122 AA.
AC Q90DN6;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UG;
RA Hu D.U., Bages J., Downing R.G., Pieniazek D., Dorn J., Fridlund C.,
RA Bivabwaho B., Sempala S.D.K., Rayfield M.A., Dondero T.J., Lal R.,
RT "Similar distribution and continued predominance of HIV-1 subtypes A
and D infections in Uganda."
RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF307695; AAL08756.1; -.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0019031; C:Viral envelope; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
KW Pfam; PF00517; GP41; 1.
KW Transmembrane.

FT NON_TER 1 1
SQ SEQUENCE 122 AA; 14728 MW; EBF439193281EDFE CRC64;

Query Match 48.4%; Score 45; DB 15; Length 122;
Best Local Similarity 60.0%; Pred. No. 12;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WIDIKOFVWY 10
Db 112 WFDITQWLWY 121

RESULT 14

Q9QIW7 PRELIMINARY; PRT; 122 AA.
AC Q9QIW7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GP1ERR04;
RA Caride E., Hertogs K., Larder B., Dehertogh P., Brindeiro R.,
RA Machado E., de Sa C.A.M., Eyer W., Passioni L.F.C., Menezes J.A.,
RA Calazans A.R., Tanuri A.;
RT "Genotyping and phenotyping analysis of B and non-B Human
immunodeficiency virus type 1 subtypes from patients under HAART."
RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF165537; AAF08482.1; -.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0019031; C:Viral envelope; IEA.

DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.

FT NON_TER 1 1
SQ SEQUENCE 122 AA; 14834 MW; 13711994A52A5B33 CRC64;

Query Match 48.4%; Score 45; DB 15; Length 122;
Best Local Similarity 60.0%; Pred. No. 12;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WIDIKOFVWY 10
Db 112 WFDITQWLWY 121

RESULT 15

Q9YXR4 PRELIMINARY; PRT; 122 AA.
AC Q9YXR4;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Envelope glycoprotein immunodominant region (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KJ96BRP009;
RA Tanuri A., Swanson P.A., Devare S.G., Berro O.J., Savedra A.,
RA Costa L.J., Telles J.G., Brindeiro R., Schable C., Pieniazek D.,
RA Rayfield M.;
RT "HIV-1 subtypes among blood donors from Rio de Janeiro, Brazil."
RL Submitted (NOV-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF034039; AAC79291.1; -.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0019031; C:Viral envelope; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR000328; Env_GP41.

DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1 1
SQ SEQUENCE 122 AA; 14739 MW; 40D4789EB8C886C2 CRC64;

Query Match 48.4%; Score 45; DB 15; Length 122;
Best Local Similarity 60.0%; Pred. No. 12;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WIDIKOFVWY 10
Db 112 WFDITQWLWY 121

Search completed: July 12, 2004, 14:10:27
Job time : 24.1512 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 13:52:02 ; Search time 4.7093 Seconds

(without alignments)
165.853 Million cell updates/sec

Title: US-10-010-667A-2_COPY_271_291

Perfect score: 93

Sequence: 1 WIDIKQFWYTPPTF 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	93	100.0	339	1 STEA_HUMAN	Q9UHE8 homo sapien
2	52	55.9	458	1 AGLE_RHME	O943X5 rhizobium m
3	45	48.4	865	1 ENV_HYLRH	P04579 human immu
4	43.5	46.8	1167	1 WCI_NEUCR	O01371 neurospora
5	43	46.2	792	1 OSTR_XYIFR	O87A19 xyella fas
6	42	45.2	346	1 OBP_ICTPU	O42266 ictalurus p
7	42	45.2	396	1 YOR1_YEAST	P45992 saccharomyc
8	42	45.2	685	1 ERP2_YEAST	P05631 saccharomyc
9	42	45.2	795	1 TUR1_MOUSE	O96DQ1 mus musculu
10	42	45.2	847	1 ENV_HY1S1	P15550 human immu
11	41.5	44.6	224	1 YN85_YEAST	P53721 saccharomyc
12	41	44.1	405	1 WCAD_ECOLI	P71238 escherichia
13	41	44.1	430	1 PUCK_BACUV	O31140 bacillus su
14	41	44.1	454	1 DHE4_SACBA	O87FF6 saccharomyc
15	41	44.1	703	1 ARYB_MANSE	P14297 marduca sex
16	41	44.1	843	1 ENV_HY1Y2	P35961 human immu
17	41	44.1	854	1 ENV_STVCZ	P11281 citampazee
18	41	44.1	856	1 ENV_HY1ZH	P05881 human immu
19	41	44.1	1509	1 MYSN_AQACA	P05659 acanthamoeb
20	41	44.1	1628	1 NAGH_CLOPE	P26831 clostridium
21	40	43.0	226	1 HA2P_RABIT	P20755 cryotlagus
22	40	43.0	792	1 OSTR_XYIFR	O9F641 xyella fas
23	40	43.0	847	1 ENV_HY1W2	P05880 human immu
24	40	43.0	855	1 ENV_HY1Z2	P04587 human immu
25	40	43.0	855	1 ENV_HY1Z6	P04587 human immu
26	40	43.0	856	1 ENV_HY1NM	P05877 human immu
27	39.5	42.5	263	1 MPPI_LOLPR	P14946 lolium pere
28	39.5	42.5	263	1 MPPI_PHLPR	P44213 phleum prat
29	39.5	42.5	265	1 MPH1_HOILA	O41216 holcus lana
30	39.5	42.5	269	1 MPH1_PHAPO	O41260 palatis ag
31	39.5	42.5	551	1 YABN_ECOLI	P33595 escherichia
32	39.5	42.5	701	1 ACSA_HUMAN	O9nr19 homo sapien
33	39	41.9	210	1 SSH5_YEAST	O03446 saccharomyc

34	39	41.9	223	1 CDR1_HUMAN	P51861 homo sapien
35	39	41.9	282	1 Y765_HAETN	O57125 haemophilus
36	39	41.9	289	1 MTM3_METUA	O58015 methanococc
37	39	41.9	361	1 SERC_BACHD	O94kM4 bacillus ha
38	39	41.9	459	1 ND4M_MACRO	P92658 macropus ro
39	39	41.9	847	1 OPGH_ECOLI6	O8f853 escherichia
40	39	41.9	847	1 OPGH_ECOLI	P33137 escherichia
41	39	41.9	847	1 OPGH_SALTI	O82718 salmoneilla
42	39	41.9	847	1 OPGH_SALTY	O82q26 salmoneilla
43	39	41.9	853	1 ENV_HY1EL	P04581 human immu
44	39	41.9	867	1 OPGH_NITRU	O82a58 nitrospora
45	39	41.9	868	1 ENV_HY1C4	P05679 human immu

ALIGNMENTS

```
RESULT 1
STEA_HUMAN
ID STEA_HUMAN STANDARD; FR; 339 AA.
AC Q9UHE8; O95034;
DR 16-OCT-2001 (Rel. 40, Created)
DR 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Six transmembrane epithelial antigen of prostate.
GN STEAP OR STEAP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20056277; PubMed=10586738;
RA Hubert R.S., Vivanco I., Chen E., Rastegar S., Leong K.,
RA Mitchell S.C., Madraswala R., Zhou Y., Kuo J., Rattano A.B.,
RA Jakobovits A., Saffran D.C., Afar D.R.H., antigen highly expressed in
RT "STEAP: a prostate-specific cell-surface antigen highly expressed in
RT human prostate tumors."
RT Proc. Natl. Acad. Sci. U.S.A. 96:14523-14528(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Abu-Threideh J., Stoneking T., Langston Y., Maupin R.,
RA Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedlin T.B., Toschinski S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Halyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Heltón E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman D.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schercher A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -!- TISSUE SPECIFICITY: Highly expressed in prostate tumors.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF186249; AAF17479.1; -.
DR EMBL; AC005053; AAC79150.1; ALT_INIT.
DR EMBL; AC004969; AAD15620.2; -.
DR EMBL; BC011802; AAH1802.1; -.
DR Genem; HGNC:11378; STEAP.
DR MIM; 604415; -.
DR GO; GO:0005887; C:intracel to plasma membrane; TAS.
DR GO; GO:0005911; C:intracellular junction; TAS.
DR GO; GO:0015267; F:channel/pore class transporter activity; TAS.
KM Transmembrane; Antigen.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 119 139 POTENTIAL.
FT TRANSMEM 164 184 POTENTIAL.
FT TRANSMEM 218 238 POTENTIAL.
FT TRANSMEM 258 278 POTENTIAL.
FT TRANSMEM 291 311 POTENTIAL.
SQ SEQUENCE 339 AA; 39851 MW; 55443A170C870387 CRC64;

Query Match 100.0%; Score 93; DB 1; Length 339;
Best Local Similarity 100.0%; Pred. No. 7.4e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WDIKQFWYTPPTF 15
DB 277 WDIKQFWYTPPTF 291
ID AGLE RHIME STANDARD; PRT; 458 AA.
AC Q923R5;
DT 30-MAY-2000 (Rel. 39, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alpha-glucosides-binding periplasmic protein aglE precursor.
GN AGLE OR R00695 OR SMC03061.
OS Rhizobium melioli (Sinorhizobium melioli).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99328961; PubMed=10400573;
RA Willis L.B., Walker G.C.;
RT "A novel Sinorhizobium melioli operon encodes an alpha-glucosidase
RT and a periplasmic-binding-protein-dependent transport system for
RT alpha-glucosides.";
RL J. Bacteriol. 181:4176-4184(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gozzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godite T., Goffeau A., Kahn D., Kiss B., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galbert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium melioli strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC FOR ALPHA-GLUCOSIDES SUCH AS SUCROSE, MALTOSE AND TREHALOSE.
CC -!- SUBCELLULAR LOCATION: Periplasmic (Probable).
CC -!- SIMILARITY: Belongs to the bacterial extracellular solute-binding
CC protein family 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF045609; AAD12050.1; ALT_INIT.
DR EMBL; AL591784; CAC45267.1; -.
DR InterPro; IPR006059; SBP_dac_1.
DR InterPro; IPR006061; SBP_domi.
DR Pfam; PF01547; SBP_dac_1; 1.
DR PROSITE; PS01037; SBP_BACTERIAL_1; FALSE NEG.
KM Sugar transporter; Transport; Periplasmic; Signal; Complete proteome.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 458 ALPHA-GLUCOSIDES-BINDING PERIPLASMIC
FT PROTEIN AGLE.
SQ SEQUENCE 458 AA; 49703 MW; E7A7F8157C2FC291 CRC64;

Query Match 55.9%; Score 52; DB 1; Length 458;
Best Local Similarity 57.1%; Pred. No. 0.5;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 IDIKQFWYTPPTF 15
DB 173 IDIKSLVWYVPENF 186
ID ENV HVLRH STANDARD; PRT; 865 AA.
AC P04579;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contactus: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (RF/HAT isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11701;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86218077; PubMed=2423250;
RA Starchic B.R., Hahn B.H., Shaw G.M., McNeely P.D., Modrow S., Wolf H.,
RA Parks E.S., Parks W.P., Joseph S.F., Gallo R.C., Wong-Staal F.;
RT "Identification and characterization of conserved and variable
RT regions in the envelope gene of HTLV-III/LAV, the retrovirus of
RT AIDS.";
RL Cell 45:637-648(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gozzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godite T., Goffeau A., Kahn D., Kiss B., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galbert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium melioli strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC FOR ALPHA-GLUCOSIDES SUCH AS SUCROSE, MALTOSE AND TREHALOSE.
CC -!- SUBCELLULAR LOCATION: Periplasmic (Probable).
CC -!- SIMILARITY: Belongs to the bacterial extracellular solute-binding
CC protein family 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```
FT DISULFID 231 260 BY SIMILARITY.
FT DISULFID 241 252 BY SIMILARITY.
FT DISULFID 309 343 BY SIMILARITY.
FT DISULFID 369 452 BY SIMILARITY.
FT DISULFID 396 425 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 193 193 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 247 247 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 254 254 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 275 275 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 344 344 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 351 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 367 367 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 407 407 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 413 413 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 455 455 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 471 471 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 634 634 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 646 646 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 825 825 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 865 AA; 97809 MM; 288286C2314AD0AC CRC64;

Query Match 48.4%; Score 45; DB 1; Length 865;
Best Local Similarity 60.0%; Pred. No. 13;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WIDKQFWY 10
Db 681 WEDITQWLMY 690

RESULT 4
WCL_NEUCR STANDARD; PRT; 1167 AA.
ID WCL_NEUCR
AC Q01371;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE White collar 1 protein (WCL).
GN WC-1.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OC NCBI_Taxid=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A / FGSC 987;
RX MEDLINE=96203083; PubMed=8612589;
RA Ballario P., Vittorioso P., Magrelli A., Talora C., Cabibbo A.,
RT "White collar-1, a central regulator of blue light responses in
RL Neurospora, is a zinc finger protein.",
RN [2]
RP EMBL J. 15:1650-1657(1996).
RP REVISIONS TO C-TERMINUS.
```

```
RA Ballario P.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR INVOLVED IN LIGHT
CC REGULATION. BINDS AND AFFECTS BLUE LIGHT REGULATION OF THE AL-3
CC GENE. WCL1 AND WCL2 PROTEINS INTERACT VIA HOMOLOGOUS PAS DOMAINS,
CC BIND TO PROMOTERS OF LIGHT REGULATED GENES SUCH AS FRQ, AND
CC ACTIVATE TRANSCRIPTION.
CC -!- SUBUNIT: HETERODIMER OF WCL1 AND WCL2 (POTENTIAL).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- INDUCTION: By blue light.
CC -!- DOMAIN: The glutamine-rich domain might function in activating
CC gene expression.
CC -!- SIMILARITY: Contains 1 GATA-type zinc finger.
CC -!- SIMILARITY: Contains 3 PAS (PBR-ARNT-SIM) dimerization domains.
CC -!- SIMILARITY: Contains 2 PAS-associated C-terminal (PAC) domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X94300; CAA63964.2; -.
DR HSRP; P17679; IGNP.
DR TRANSFAC; T02819; -.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS domain.
DR InterPro; IPR00679; Znf_GATA.
DR Pfam; PF00320; GATA; 1.
DR Pfam; PF00785; PAC; 1.
DR Pfam; PF00989; PAS; 2.
DR SMART; SM00086; PAC; 2.
DR SMART; SM00091; PAS; 3.
DR SMART; SM00401; ZNF_GATA; 1.
DR TIGRFAMs; TIGR00229; sensory_box; 3.
DR PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
DR PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
DR PROSITE; PS50112; PAS; 3.
DR KW Transcription regulation; Activator; DNA-binding; Zinc-finger;
KW Nuclear protein; Repeat.
FT FT DOMAIN 16 61 GLN-RICH.
FT FT DOMAIN 381 452 PAS 1.
FT FT DOMAIN 469 508 PAC 1.
FT FT DOMAIN 574 644 PAS 2.
FT FT DOMAIN 650 691 PAC 2.
FT FT DOMAIN 693 763 PAS 3.
FT FT ZN_FINGER 934 959 GATA-TYPE.
FT FT DOMAIN 21 57 POLY-GLN.
FT FT DOMAIN 329 333 POLY-PRO.
SQ SEQUENCE 1167 AA; 127454 MM; 6489D04DAB50EE38 CRC64;

Query Match 46.8%; Score 43.5; DB 1; Length 1167;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 8; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 3 DIKQFWYTPPT 14
Db 529 DIQYIM-TTPPT 539

RESULT 5
OSTA_XYLFT STANDARD; PRT; 792 AA.
ID OSTA_XYLFT
AC Q87A19;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Organic solvent tolerance protein precursor.
GN IMP OR OSTA OR PD1836.
OS Xylella fastidiosa (strain Temecual / ATCC 700964).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
```

```
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=183190;
RN [1]
RX SEQUENCE FROM N.A.
RA MEDLINE=2242131; PubMed=12533478;
RA Van Slyus M.A., de Oliveira M.C., Monteiro-Vitorello C.B.,
RA Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., Moon D.H.,
RA Takita M.A., Lemos E.G.M., Machado M.A., Ferro M.I.T., da Silva F.R.,
RA Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorri H., Tsai S.M.,
RA Carrier H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira M.U.,
RA Coutinho L.L., Kimura E.T., Ferro E.S., Harakava R., Kuramae E.E.,
RA Maingot C.L., Gigliotti E., Abreu I.L., Alves L.M.C., do Amaral A.V.,
RA Bairo G.S., Blanco S.R., Brito M.S., Cannavan F.S., Celestino A.V.,
RA da Cunha A.F., Fennille R.C., Ferro J.A., Formighieri E.F., Kishi L.T.,
RA Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Saesaki F.T., Sena J.A.D.,
RA de Souza A.A., Trifil D., Tsukumo F., Yanai G.M., Zares L.G.,
RA Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
RA Kitajima J.P.;
RT "Comparative analyses of the complete genome sequences of Pierce's
RT disease and citrus variegated chlorosis strains of Xylella
RT fastidiosa.";
RT J. Bacteriol. 185:1018-1026(2003).
CC -1- FUNCTION: Determines N-hexane tolerance. Involved in outer
CC membrane permeability. Essential for envelope biogenesis. Could be
CC part of a targeting/usher system for outer membrane components (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Outer membrane (By similarity).
CC -1- SIMILARITY: Belongs to the Imp/OscA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AE012560; AAO29668.1; -.
DR HAMAP; MF_01411; 1.
DR InterPro; IPR007543; OSCA_C.
DR Pfam; PF04453; OSCA_C; 1.
DR Kew Outer membrane; Signal; Complete proteome.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 792 ORGANIC SOLVENT TOLERANCE PROTEIN.
SQ SEQUENCE 792 AA; 90768 MW; 70637D6FA7B62DAD CRC64;
Query Match 46.2%; Score 43; DB 1; Length 792;
Best Local Similarity 36.4%; Pred. No. 26;
Matches 8; Conservative 3; Mismatches 3; Indels 8; Gaps 1;
QY 2 IDIKQFY-----WYTPPTP 15
Db 440 IDVKPYSLPFTGASWYTPPTP 461
RESULT 6
OPSP ICTPU STANDARD; PRT; 346 AA.
ID OPSP ICTPU
AC 04226;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Parapinopsin.
OS Ictalurus punctatus (Channel catfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
OC Ictaluridae; Ictalurus.
OX NCBI_TaxID=7998;
RN [1]
RX SEQUENCE FROM N.A.
RA MEDLINE=97477428; PubMed=9334384;
RA Blackshaw S., Snyder S.H.;
RT "Parapinopsin, a novel catfish opsin localized to the paraneal
```

```
RT organ, defines a new gene family.";
RL J. Neurosci. 17:8083-8092(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: PARAPINEAL ORGAN.
CC -1- PTM: Some or all of the carboxyl-terminal Ser or Thr residues may
CC be phosphorylated (By similarity).
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC Opisin subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF028014; AAB84050.1; -.
DR HSSP; P02699; 1FDP.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR001760; Opisin.
DR Pfam; PF00001; 7cm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSN.
DR PROSITE; PS00237; G PROTEIN RECEPT_F1_1; 1.
DR PROSITE; PS00262; G PROTEIN RECEPT_F1_2; 1.
DR PROSITE; PS00238; OESIN; 1.
KW Photoreceptor; Retinal protein; Transmembrane; Glycoprotein;
KW Phosphorylation; Lipoprotein; Palmitate; G-protein coupled receptor.
FT DOMAIN 1 29
FT TRANSMEM 30 54
FT DOMAIN 55 66
FT TRANSMEM 67 91
FT DOMAIN 92 106
FT TRANSMEM 107 126
FT DOMAIN 127 145
FT TRANSMEM 146 169
FT DOMAIN 170 193
FT TRANSMEM 194 221
FT DOMAIN 222 244
FT TRANSMEM 245 268
FT TRANSMEM 269 276
FT TRANSMEM 277 301
FT DOMAIN 302 346
FT DISULFID 103 180
FT BINDING 288 288
FT CARBOHYD 8 8
FT CARBOHYD 191 191
FT LIPID 315 315
SQ SEQUENCE 346 AA; 38203 MW; A70871684F8FC7FD CRC64;
Query Match 45.2%; Score 42; DB 1; Length 346;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 7 FVWYTPPTP 15
Db 158 FVWYTPPTP 166
RESULT 7
YURL YEAST STANDARD; PRT; 396 AA.
ID YURL YEAST
AC P46992;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical 43.0 kDa protein in CPS1-FP1 intergenic region.
GN YJL171C OR J0512
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
```

RP SEQUENCE FROM N.A.
RA Obermaier B., Piravandi E., Rinke M., Dordney H.;
RL Submitted (SEP-1995) to the EMBL/Genbank/DBJ databases.
CC
CC -1- SIMILARITY: TO YEAST YER162C.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Z49446; CAA89466.1; -
DR PIR: S56954; S56954.
DR Germonline: 141783; -
DR SGD: S0003707; YOL171C.
KW Hypothetical protein.
SQ SEQUENCE 396 AA; 43014 MW; 279E858E7512670A CRC64;

Query Match 45.2%; Score 42; DB 1; Length 396;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 IDIKOFVWYTPPT 15
Db 83 LKXQFARYISPGF 96

RESULT 8
ERF2_YEAST
ID ERF2_YEAST STANDARD; PRT; 685 AA.
AC P05453; P05420; (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Eukaryotic peptide chain release factor GTP-binding subunit (ERF2)
DE (Translation release factor 3) (ERF3) (ERF-2) (Omnipotent suppressor
DE protein 2) (G1 to S phase transition protein 1).
GN SUP35 OR SUP2 OR SUP12 OR GST1 OR SAL3 OR PMW2 OR YDR172W OR
GN YD9395.05.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
KN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88329727; PubMed=3047009;
RA Kushnir V.V., Ter-Avanesyan M.D., Telckov M.V., Surguchov A.P.,
RT Smirnov V.N., Inge-Vechtomov S.G.;
RT "Nucleotide sequence of the SUP2 (SUP35) gene of Saccharomyces
RT cerevisiae.";
RL Gene 66:45-54 (1988).
RL [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87219095; PubMed=3556215;
RA Inge-Vechtomov S.G.,
RA Kushnir V.V., Ter-Avanesyan M.D., Surguchov A.P., Smirnov V.N.,
RT "Localization of possible functional domains in sup2 gene product of
RT the yeast Saccharomyces cerevisiae.";
RL FEBS Lett. 215:257-260 (1987).
RL [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88172503; PubMed=3280807;
RA Wilson P.G., Culbertson M.R.;
RT "SUP12 suppressor protein of yeast. A fusion protein related to the
RT EF-1 family of elongation factors.";
RL J. Mol. Biol. 199:559-573 (1986).
RL [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=88296422; PubMed=2841115;
RA Kikuchi Y., Shimatake H., Kikuchi A.;
RT "A yeast gene required for the G1-to-S transition encodes a protein

RT containing an A-kinase target site and GTPase domain.";
RL EMBL J. 7:1175-1182 (1986).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Murphy L., Harris D.E., Barrell B.G., Rajandream M.A.;
RL Submitted (NOV-1994) to the EMBL/Genbank/DBJ databases.
RN [6]
RP FUNCTION.
RX MEDLINE=96016209; PubMed=7556078;
RA Stansfield I., Jones K.M., Kushnir V.V., Dagkesamanskaya A.R.,
RA Poznyakovsky A.I., Pauskin S.V., Nierres C.R., Cox B.S.,
RA Ter-Avanesyan M.D., Tulce M.F.;
RT "The products of the SUP45 (erf1) and SUP35 genes interact to mediate
RT translation termination in Saccharomyces cerevisiae.";
RL EMBL J. 14:4365-4373 (1995).
CC -1- FUNCTION: INVOLVED IN TRANSLATION TERMINATION. STIMULATES THE
CC ACTIVITY OF ERFL. BINDS GUANINE NUCLEOTIDES.
CC -1- SUBUNIT: Heterodimer of two subunits, one of which binds GTP.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC ERF3 SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M21129; AAA35133.1; -
DR EMBL: X07163; CAA30155.1; -
DR EMBL: Y00829; CAA68760.1; -
DR EMBL: Z46727; CAA86677.1; -
DR PIR: S00733; EFBYS2.
DR Germonline: 140663; -
DR SGD: S0002579; SUP35.
DR InterPro: IPR000795; EF_GTPbind.
DR InterPro: IPR004160; EFTU Cterm.
DR InterPro: IPR004161; EFTU D2.
DR InterPro: IPR009001; Elongat_init C.
DR InterPro: IPR009000; Translat_factor.
DR Pfam: PF00009; GTP_EFTU; 1.
DR Pfam: PF03144; GTP_EFTU_D2; 1.
DR Pfam: PF03143; GTP_EFTU_D3; 1.
DR PRINTS: PR00315; ELONGATN_FCT.
DR PROSITE: PS00301; EFACITOR_GTP; 1.
KW Protein biosynthesis; Repeat; Phosphorylation; GTP-binding.
FT DOMAIN 5 135
FT DOWAIN 5 135
FT DOWAIN 5 135
FT NP_BIND 267 249
FT NP_BIND 267 249
FT NP_BIND 344 348
FT NP_BIND 344 348
FT NP_BIND 406 409
FT NP_BIND 406 409
FT ACT_SITE 273 273
FT ACT_SITE 273 273
FT ACT_SITE 407 407
FT MOD_RES 341 341
FT MOD_RES 341 341
FT CONFLICT 53 53
SQ SEQUENCE 685 AA; 76551 MW; 43912A6D7DFA153 CRC64;

Query Match 45.2%; Score 42; DB 1; Length 685;
Best Local Similarity 53.8%; Pred. No. 32;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 IDIKOFVWYTPPT 14
Db 459 VDPKCEPVTGTPT 471

RESULT 9
TIR1_MOUSE
ID TIR1_MOUSE STANDARD; PRT; 795 AA.
AC Q9EPQ1; Q9EPW5;

DR DT 28-FEB-2003 (Rel. 41, Created)
DT DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Toll-like receptor 1 precursor (Toll/interleukin-1 receptor-like)
DE (TIR).
GN TLR1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid:10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Macrophage;
RX MEDLINE=20558581; PubMed=11095749;
RA Ozinsky A., Underhill D.M., Fontenot J.D., Hajjar A.M., Smith K.D.,
RA Wilson C.B., Schroeder L., Aderem A.;
RT "The repertoire for pattern recognition of pathogens by the innate
RT immune system is defined by cooperation between Toll-like
RT receptors.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:13766-13771 (2000).
RL [2]
RN SEQUENCE FROM N.A.
RP TISSUE=Macrophage;
RC MEDLINE=20571875; PubMed=11123271;
RX Hajjar A.M., O'Mahony D.S., Ozinsky A., Underhill D.M., Aderem A.,
RA Klebanoff S.J., Wilson C.B.;
RT "Functional interactions between Toll-like receptor (TLR) 2 and TLR1
RT or TLR6 in response to phenol-soluble modulin.";
RL J. Immunol. 166:15-19 (2001).
RL [3]
RN SEQUENCE FROM N.A.
RP STRAIN=BALB/c; TISSUE=Spleen;
RC Thomson D.P., Campbell C.C., Liew F.Y., Xu D.;
RT "Cloning of Mus musculus Toll-like receptor 1.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Participates in the innate immune response to microbial
CC agents. Cooperates with TLR2 and modulates the response to
CC microbial constituents. Acts via MyD88 and TRAF6, leading to NF-
CC kappa-B activation, cytokine secretion and the inflammatory
CC response (By similarity).
CC -1- SUBUNIT: Binds TLR2 via their respective extracellular domains.
CC Binds MyD88 via their respective TIR domains (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Plasma membrane and
CC phagosomes.
CC -1- SIMILARITY: Belongs to the Toll-like receptor family.
CC -1- SIMILARITY: Contains 1 TIR domain.
CC -1- SIMILARITY: Contains 9 leucine-rich (LRR) repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC -----
DR EMBL; AY009154; AAC37302.1; -
DR EMBL; AF316985; AAG35062.1; -
DR HSSP; O60603; IPYM.
DR MGD; MGI:1341295; TIR1.
DR GO; GO:0016020; C:membrane; NAS.
DR GO; GO:0045335; C:phagocytic vesicle; NAS.
DR GO; GO:0003793; F:defense/immunity protein activity; NAS.
DR GO; GO:0004888; F:transmembrane receptor activity; NAS.
DR GO; GO:0002497; F:triacylated lipopeptide binding; NAS.
DR GO; GO:0007250; P:activation of NF-kappa-B-inducing kinase; NAS.
DR GO; GO:0042116; P:macrophage activation; NAS.
DR GO; GO:0042945; P:perception of triacylated bacterial lipopo. . .; NAS.
DR GO; GO:0045410; P:positive regulation of interleukin-6 biosyn. . .; IMP.
DR GO; GO:0042355; P:positive regulation of tumor necrosis facto. . .; IMP.
DR InterPro; IPR004075; IIL_receptor1.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.

DR	InterPro; IPR000157; TIR.
DR	Pfam; PF00560; LRR; 8.
DR	Pfam; PF01463; LRRT; 1.
DR	Pfam; PF01582; TIR; 1.
DR	PRINTS; PR01537; INTRLNRI.F.
DR	PRINTS; PR00019; LEURICHRPT.
DR	SMART; SMO0082; LRCT; 1.
DR	SMART; SMO0255; TIR; 1.
DR	PROSITE; PS50104; TIR; 1.
KW	Receptor; Immune response; Inflammatory response; Signal;
KW	Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein.
FT	SIGNAL
FT	CHAIN
FT	DOMAIN
FT	TRANSMEM
FT	DOMAIN
FT	REPEAT
FT	REPEAT
FT	REPEAT
FT	REPEAT
FT	REPEAT
FT	REPEAT
FT	REPEAT
FT	REPEAT
FT	DOMAIN
FT	CARBOHYD
FT	CARBOHYD
FT	CARBOHYD
FT	CARBOHYD
FT	CARBOHYD
FT	CARBOHYD
FT	CARBOHYD
FT	CARBOHYD
FT	CARBOHYD
FT	CONFLICT
SO	SEQUENCE
Qy	1 WIDKGFVWTTP 12
Dd	264 FINTLIQIVMTTP 275
ENV_HYV1S1	STANDARD; PRT; 847 AA.
AC	P19550;
DT	01-FEB-1991 (Rel. 17, Created)
DT	01-FEB-1991 (Rel. 17, Last sequence update)
DT	10-OCT-2003 (Rel. 42, last annotation update)
DE	Envelope polypeptide GP160 precursor [contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
DN	ENV.
GN	Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).
OC	Vitruvius; Retrovirus; Retroviridae; Lentiviruses.
OX	NCBI_Taxid=11691;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=90347835; PubMed=2384920;
RA	Cheng-Mayer C., Quiroga M., Tung J.W., Dina D., Levy J.
RT	"Viral determinants of human immunodeficiency virus type 1 T-cell or macrophage tropism, cytopathogenicity, and CD4 antigen modulation." ;
RL	J. Virol. 64:4390-4398(1990).
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; M65024; AAA45072.1; -.
DR PDB; 10BE; 15-MAY-97.
DR HIV; M38428; ENVSEF162.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00517; GP41; 1.
DR Pfam; PF00517; GP120; 1.
DR Aids; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
  3D-structure.
KW SIGNAL.
FT CHAIN 1 29
FT CHAIN 30 502
FT CHAIN 503 847
FT DISULFID 53 73
FT DISULFID 118 203
FT DISULFID 125 194
FT DISULFID 130 155
FT DISULFID 216 245
FT DISULFID 226 237
FT DISULFID 294 328
FT DISULFID 374 435
FT DISULFID 381 408
FT CARBOHYD 87 87
FT CARBOHYD 135 135
FT CARBOHYD 154 154
FT CARBOHYD 186 186
FT CARBOHYD 195 195
FT CARBOHYD 232 232
FT CARBOHYD 239 239
FT CARBOHYD 260 260
FT CARBOHYD 274 274
FT CARBOHYD 293 293
FT CARBOHYD 299 299
FT CARBOHYD 329 329
FT CARBOHYD 336 336
FT CARBOHYD 352 352
FT CARBOHYD 382 382
FT CARBOHYD 388 388
FT CARBOHYD 392 392
FT CARBOHYD 398 398
FT CARBOHYD 401 401
FT CARBOHYD 438 438
FT CARBOHYD 454 454
FT CARBOHYD 602 602
FT CARBOHYD 607 607
FT CARBOHYD 616 616
FT CARBOHYD 628 628
SQ SEQUENCE 847 AA; 96135 MW; 0A901317PD7FF2AB CRC64;

Query Match 45.2%; Score 42; DB 1; Length 847;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 WIDIKOFVWYPT 10
Db 663 WFDISKMLWY 672

RESULT 11
YN89 YEAST STANDARD; PRT; 224 AA.
AC P53721;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical 25.3 kDa protein in TIM3-ARE2 intergenic region.
GN YN018W OR N1815.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
NCBI_TaxID=4932;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RA Pohl T.M.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; Z71633; CAA96297.1; -.
DR PIR; S63349; S63349.
DR GeronOnline; 143363; -.
DR SCD; S0005301; YNR018W.
DR InterPro; IPR007667; HIG_1_N.
DR Pfam; PF04588; HIG_1_N; 1.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 19 39
FT TRANSMEM 50 70
FT TRANSMEM 152 172
FT DOMAIN 196 199
SQ SEQUENCE 224 AA; 25344 MW; FA2C528A008CE7C CRC64;

Query Match 44.6%; Score 41.5; DB 1; Length 224;
Best Local Similarity 57.1%; Pred. No. 12;
Matches 8; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 1 WIDIKOFVWYPTPT 14
Db 52 W-SIKTALWIRPPT 64

RESULT 12
WCAD ECOLI STANDARD; PRT; 405 AA.
AC P71238; P76385;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative colanic acid polymerase.
GN WCAD OR B2056.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
NCBI_TaxID=562;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=96326333; PubMed=8759852;
RA Stevenson G., Andrianopoulos K., Hobbs M., Reeves P.R.;
RT "Organization of the Escherichia coli K-12 gene cluster responsible
for production of the extracellular polysaccharide colanic acid.";
RL J. Bacteriol. 178:4885-4893(1996).
[2]
RP REVISIONS.
RC STRAIN=K12;
RA Reeves P.R.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blotner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
Science 277:1453-1474(1997).
[4]
RP -1- PATHWAY: Slime polysaccharide colanic acid biosynthesis.
```



```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U38473; AAC77839.1; -.
CC EMBL: AB000295; AAC75117.1; -.
CC PIR: G64971; G64971.
CC Ecogene: EG13572; wcad.
CC LipoPolysaccharide biosynthesis; Complete proteome.
CC CONFLICT 108 108 P>A (IN REF. 2).
CC SEQUENCE 405 AA; 45409 MW; A3D9D91255686043 CRC64;
SQ
Query Match 44.1%; Score 41; DB 1; Length 405;
Best Local Similarity 57.1%; Pred. No. 27;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 WDIKQFVWYTPPT 14
Db 188 WLSIKQFGIKTPT 201
RESULT 13
PUCB_BACSU STANDARD; PRT; 430 AA.
AC 032140;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Uric acid permease pucb.
GN PUCB OR BS032440.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
ON NCBI_TaxID=1423;
RX STRAIN=168;
RA MEDLINE=9804403; PubMed=9384377;
RA Kunze F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Brun M., Bignon S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.K.,
RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
RA Denzot F., Devigne K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Eutzan K.D., Erington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golligly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koelter P., Konigstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetlelle D., Porwollik S., Prescott A.M.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha F., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serron P., Shin B.S., Soldo B.,
RA Sorokin A., Taccoti E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandepol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winers P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zuchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
```

```
RN [2]
RP FUNCTION.
RX STRAIN=168;
RA MEDLINE=21242727; PubMed=11344136;
RA Schults A.C., Mygaard P., Saxild H.H.;
RT "Functional analysis of 14 genes that constitute the purine catabolic
RT pathway in Bacillus subtilis and evidence for a novel regulon
RT controlled by the PucR transcription activator."
RL J. Bacteriol. 183:3293-3302(2001).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (probable).
CC -1- INDUCTION: Expression is very low in excess nitrogen (glutamate
CC plus ammonia) and is induced during limiting nitrogen conditions
CC (glutamate). Expression is further induced when allantoin or uric
CC acid are added during limiting nitrogen conditions.
CC -1- SIMILARITY: BELONGS TO THE XANTHINE/URACIL PERMEASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z99120; CAB15234.1; -.
CC PIR: F70016; F70016.
CC Subtilist; BG13985; pucb.
CC InterPro: IPR006042; Xan_ur_permease.
CC DR Pfam: PF00860; xan_ur_permease; 1.
CC DR TIGRFAMs: TIGR00801; ncs2; 1.
CC DR PROSITE: PS01116; XANTH URACIL PERMEASE; 1.
CC KM Transmembrane; Transport; Complete proteome.
CC FT TRANSMEM 18 38 POTENTIAL.
CC FT TRANSMEM 43 63 POTENTIAL.
CC FT TRANSMEM 67 87 POTENTIAL.
CC FT TRANSMEM 97 117 POTENTIAL.
CC FT TRANSMEM 122 142 POTENTIAL.
CC FT TRANSMEM 153 183 POTENTIAL.
CC FT TRANSMEM 185 205 POTENTIAL.
CC FT TRANSMEM 209 229 POTENTIAL.
CC FT TRANSMEM 233 253 POTENTIAL.
CC FT TRANSMEM 274 294 POTENTIAL.
CC FT TRANSMEM 310 330 POTENTIAL.
CC FT TRANSMEM 333 353 POTENTIAL.
CC FT TRANSMEM 369 389 POTENTIAL.
CC FT TRANSMEM 398 418 POTENTIAL.
CC SQ SEQUENCE 430 AA; 44861 MW; 9FD18D100A0572F CRC64;
Query Match 44.1%; Score 41; DB 1; Length 430;
Best Local Similarity 33.3%; Pred. No. 29;
Matches 5; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
QY 1 WDIKQFVWYTPPT 15
Db 215 WLVHPSLFFGPTF 229
RESULT 14
ID DHE4_SACBA STANDARD; PRT; 454 AA.
AC 08TF66;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE NADP-specific glutamate dehydrogenase 1 (EC 1.4.1.4) (NADP-GDH 1).
GN GDH1.
OS Saccharomyces bayanus (Yeast) (Saccharomyces uvarum).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
ON NCBI_TaxID=4931;
RX [1]
```



```

RP SEQUENCE FROM N.A.
RC STRAIN=CLIB 533 CBS 7001 derivative;
RA "New S. pastorianus strains and Saccharomyces natural hybrids revealed
RT by polyphasic identification of CBS strains formerly uncompletely
RT identified by conventional method."
RU Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -! CAZMLYTIC ACTIVITY: L-glutamate + H(2)O + NADP(+) = 2-oxoglutarate
CC + NH(3) + NADPH.
CC -! SUBUNIT: Homohexamer (By similarity).
CC -! SIMILARITY: Belongs to the Glu/Leu/Phe/Val dehydrogenases family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ418037; CAD10750.1; -.
DR HSSP; P96110; 1B26.
DR InterPro; IPR006095; GLFV_dehydrog.
DR InterPro; IPR006096; GLFV_dehydrog_C.
DR InterPro; IPR006097; GLFV_dehydrog_N.
DR Pfam; PF00208; GLFV_dehydrog; 1.
DR Pfam; PF02812; GLFV_dehydrog_N; 1.
DR PRINTS; PR00082; GLFDHGNASE.
DR PROSITE; PS00074; GLFV_DEHYDROGENASE; 1.
DR Oxidoreductase; NADP.
FM ACT SITE 110 110 BY SIMILARITY.
FT NP BIND 174 203 NAD (BY SIMILARITY).
SQ SEQUENCE 454 AA; 49425 MW; 98DF52A68BDEDD12 CRC64;

Query Match 44.1%; Score 41; DB 1; Length 454;
Best Local Similarity 60.0%; Pred. No. 31;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 IKQFWYTPP 13
   |||
Db 361 LKESVWYGP 370

RESULT 15
ARYB MANSE STANDARD; PRT; 703 AA.
AC P14297;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Arylphorin beta subunit precursor.
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Spingioidea;
OC Spingidae; Spinginae; Manduca.
OX NCBI_TaxID=7130;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Larval fat body;
RX MEDLINE=90037032; PubMed=2808410;
RA Willott E., Wang X.-Y., Wells M.A.;
RT "cDNA and gene sequence of Manduca sexta arylphorin, an aromatic
RT amino acid-rich larval serum protein. Homology to arthropod
RT hemocyanins."
RU J. Biol. Chem. 264:19052-19059(1989).
CC -! FUNCTION: Arylphorin is a larval storage protein (LSP) which may
CC serve as a storage protein used primarily as a source of aromatic
CC amino acids for protein synthesis during metamorphosis. It is a
CC constituent of the sclerotizing system of the cuticle, and serves
CC as a carrier for ecdysteroid hormone.
CC -! SUBUNIT: Arylphorin is a hexamer of subunits alpha and beta.
CC -! SUBCELLULAR LOCATION: Extracellular.
CC -! TISSUE SPECIFICITY: Fat body.

```

```

CC -! SIMILARITY: TO ARYB, TO B.MORI STORAGE PROTEINS 1 AND 2, AND TO
CC ARTHROPOD HEMOCYANINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M28395; AAA29304.1; -.
DR EMBL; M28397; AAA29305.1; -.
DR PIR; B34434; B34434.
DR HSSP; P04253; 10XY.
DR InterPro; IPR008922; Di-copper centre.
DR InterPro; IPR000896; Hemocyanin.
DR InterPro; IPR005203; hemocyanin_C.
DR InterPro; IPR005204; hemocyanin_N.
DR InterPro; IPR007110; 1g-1like.
DR Pfam; PF00372; hemocyanin; 1.
DR Pfam; PF03723; hemocyanin_C; 1.
DR Pfam; PF03722; hemocyanin_N; 1.
DR PRINTS; PR00187; HAEMOCYANIN.
DR PROSITE; PS00209; HEMOCYANIN_1; 1.
DR PROSITE; PS00210; HEMOCYANIN_2; 1.
DR Signal; Storage protein; Glycoprotein; Multigene family.
FM SIGNAL 1 16
FT CHAIN 17 703 ARYLPHORIN BETA SUBUNIT.
FT CARBOHYD 72 72 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 211 211 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 703 AA; 83848 MW; 24B7DBBB60D2R3FA CRC64;

Query Match 44.1%; Score 41; DB 1; Length 703;
Best Local Similarity 60.0%; Pred. No. 48;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 DIKQFWYTP 12
   |||
Db 287 DIPERSWSP 296

```

Search completed: July 12, 2004, 14:08:04
 Job time : 5.7093 secs

This Page Blank (uspto)

Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:01:53 ; Search time 7.5 Seconds

(without alignments)
192.383 Million cell updates/sec

Title: US-10-010-667a-2_COPY_271_291

Perfect score: 93

Sequence: 1 WIDIKQFWYTPPTF 15

Scoring table: BLOSUM62

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	51.6	453	2	A97431
2	48	51.6	453	2	AB2649
3	45	48.4	31	2	A23341
4	45	48.4	358	2	S21998
5	44.5	47.8	553	2	AC0065
6	44.5	47.8	662	2	E69274
7	44	47.3	331	2	AE3544
8	44	47.3	361	2	AH1469
9	44	47.3	425	2	T27241
10	44	47.3	568	2	A89958
11	43.5	46.8	1154	2	S69206
12	43	46.2	402	2	T14669
13	43	46.2	418	2	T14946
14	43	46.2	863	2	AS5034
15	43	46.2	877	2	AS49197
16	42	45.2	133	2	G81710
17	42	45.2	321	2	E81411
18	42	45.2	396	2	SS6954
19	42	45.2	462	2	B86262
20	42	45.2	639	2	E72336
21	42	45.2	685	1	EBFYS2
22	41.5	44.6	126	2	G69990
23	41.5	44.6	224	2	S63349
24	41	44.1	254	2	D69140
25	41	44.1	357	2	S21992
26	41	44.1	359	2	T44816
27	41	44.1	368	2	P84300
28	41	44.1	405	2	E90986
29	41	44.1	405	2	H85831

30	41	44.1	405	2	G64971	Purative colanic a
31	41	44.1	430	2	F70016	putine permease ho
32	41	44.1	703	2	B34434	arythorin beta ch
33	41	44.1	843	1	H44001	env polyprotein pr
34	41	44.1	847	2	T09448	env polyprotein pr
35	41	44.1	847	2	S13289	env protein - huma
36	41	44.1	852	2	T12016	env polyprotein pr
37	41	44.1	854	1	VCJST	env polyprotein pr
38	41	44.1	856	1	A44963	env polyprotein pr
39	41	44.1	929	2	G81036	isolucyl-tRNA syn
40	41	44.1	929	2	B81982	probable isolucan
41	41	44.1	1042	2	S43904	hyaluronidase - Cl
42	41	44.1	1509	1	A27224	myosin heavy chain
43	40.5	43.5	313	2	A12642	flagellin [importe
44	40.5	43.5	313	2	B97425	flac protein (X964
45	40.5	43.5	1207	2	T00378	KIM0641 protein -

ALIGNMENTS

RESULT 1
A97431
alpha-glucosides-binding periplasmic protein agle precursor [imported] - Agrobacterium
C/Species: Agrobacterium tumefaciens
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C/Accession: A97431
R./Goodner, B.; Hinkle, G.; Gattling, S.; Miller, N.; Blanchard, M.; Ourullo, B.; Goldman
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B
Science 294, 2323-2328, 2001
A./Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu
A./Reference number: A97359; MUID:21608551; PMID:11743194
A./Accession: A97431
A./Status: preliminary
A./Molecule type: DNA
A./Residues: 1-453 <KID>
A./Cross-references: GB:AB007869; PIDN:AAK6402.1; PID:G15155534; GSPDB:GN00169
C./Genetics:
A./Gene: AGR_C 1045
A./Map position: circular chromosome

Query Match 51.6%; Score 48; DB 2; Length 453;
Best Local Similarity 53.8%; Pred. No. 4.1;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 3 DIKQFWYTPPTF 15
DB 169 DVKSLVWVPENF 181

RESULT 2
AB2649
hypothetical protein agle [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C/Accession: AB2649
R./Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavina, T.; Levy, R.; Li, M.; McClell
Science 294, 2317-2323, 2001
A./Authors: Yoo, H.; Tso, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A./Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A./Reference number: AB2577; MUID:21608550; PMID:11743193
A./Accession: AB2649
A./Status: preliminary
A./Molecule type: DNA
A./Residues: 1-453 <KID>
A./Cross-references: GB:AE008688; PIDN:AAL1608.1; PID:G17738945; GSPDB:GN00186
A./Experimental source: strain C58 (Dupont)
C./Genetics:
A./Gene: agle
A./Map position: circular chromosome

Query Match 51.6%; Score 48; DB 2; Length 453;
Best Local Similarity 53.8%; Pred. No. 4.1;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 DIKQFVWYPTPT 15
|:|:|:|:|:|:|
Db 169 DVKSLWVYVPENP 181

RESULT 3
A23341
allergen R7 - perennial ryegrass (tentative sequence) (fragment)
C:Species: Lolium perenne (perennial ryegrass)
C>Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 09-Jun-2000
C:Accession: A23341
R: Cottam, G.P.; Moran, D.M.; Standing, R.
Biochem. J. 234, 305-310, 1986
A:Title: Physicochemical and immunochemical characterization of allergenic proteins from
A:Reference number: A23341; PMID:86242068; PMID:3718469

A:Accession: A23341
A:Molecule type: protein
A:Residues: 1-31 <COT>
C:Superfamily: expansin

Query Match 48.4%; Score 45; DB 2; Length 31;
Best Local Similarity 50.0%; Pred. No. 0.65;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 WIDIKQFVWYPTPT 14
|:|:|:|:|:|:|
Db 18 WIDAKSTWYVKRPT 31

RESULT 4
S21998
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C:Species: human immunodeficiency virus type 1, HIV-1
A:Variety: isolate 28
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C:Accession: S21998; S70425
R: Steiner, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
Submitted to the EMBL Data Library, July 1991
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determi
A:Reference number: S21998
A:Accession: S21998
A:Molecule type: DNA
A:Residues: 1-358 <STET>
A:Cross-references: EMBL:X61359; NID:g60182; PIDN:CAA43630.1; PID:g60183
R: Steiner, H.; Storch-Hagenlocher, B.; Wildemann, B.
AIDS Res. Hum. Retroviruses 8, 53-59, 1992
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebr
A:Reference number: S70417; PMID:92144209; PMID:1736940
A:Accession: S70425
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-222 'X', 224-358 <STET>
A:Cross-references: EMBL:X61359; NID:g60182; PIDN:CAA43630.1; PID:g60183
C:Superfamily: type B retrovirus env polyprotein

Query Match 48.4%; Score 45; DB 2; Length 358;
Best Local Similarity 60.0%; Pred. No. 9.9;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WIDIKQFVWY 10
|:|:|:|:|:|:|
Db 174 WFDITQWY 183

RESULT 5
AC0065
conserved hypothetical protein YP00524 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis

C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C:Accession: AC0065
R: Parthill, U.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.
demo-Parraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; PMID:121470413; PMID:11586360
A:Accession: AC0065
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-553 <KUR>
A:Cross-references: GB:AL590842; PIDN:CA089382.1; PID:g15978619; GSPDB:GN00175
C:Genetics: YP00524

Query Match 47.8%; Score 44.5; DB 2; Length 553;
Best Local Similarity 53.8%; Pred. No. 20;
Matches 7; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 1 WIDIKQFVWYPTPT 13
|:|:|:|:|:|:|
Db 540 WDFKX-AWFTPT 551

RESULT 6
E69274
acetyl-CoA synthetase (acs-1) homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jan-2000
C:Accession: E69274
R: Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
., J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weisman, U.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Ullrich, T.; Cottam, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: E69250; PMID:98049343; PMID:9389475
A:Accession: E69274
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-662 <KUR>
A:Cross-references: GB:AB001092; GB:AE000782; NID:g2689415; PIDN:AAB91033.1; PID:g265044
C:Superfamily: acetate-CoA ligase; acetate-CoA ligase homology
F:145-621/Domain: acetate-CoA ligase homology <ACU>

Query Match 47.8%; Score 44.5; DB 2; Length 662;
Best Local Similarity 41.2%; Pred. No. 24;
Matches 7; Conservative 6; Mismatches 1; Indels 3; Gaps 1;

QY 2 DIKQ---FWYPTPTPT 15
|:|:|:|:|:|:|
Db 307 WDFKGDRLWYSPSW 323

RESULT 7
AF3544
Hypothetical membrane spanning protein BMEI10279 [imported] - Brucella melitensis (strai
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AF3544
R: DelVecchio, V.G.; Kaprat, V.; Redkar, R.J.; Patriz, G.; Muir, C.; Los, T.; Ivanova,
., Mazur, M.; Goltzman, E.; Selkov, E.; Elser, P.H.; Hagius, S.; O'Callaghan, D.; Letes
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A:Reference number: AD3252; PMID:11756688
A:Accession: AF3544
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-331 <KUR>
A:Cross-references: GB:AE008918; PIDN:AAU53521.1; PID:g17984427; GSPDB:GN00191
A:Experimental source: strain 16M

C:Genetics:
A:Gene: BMEI10279
A:Map position: 11

Query Match 47.3%; Score 44; DB 2; Length 331;
Best Local Similarity 38.5%; Pred. No. 13;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 DIKQFWYTPPT 15
DB 117 NVKDRVWHTPDPF 129

RESULT 8

AHL469

Internalin protein homolog lin0295 [imported] - *Listeria innocua* (strain Clij11262)

C:Species: *Listeria innocua*

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C:Accession: AHL469

R:Glaser, P.; Frauguel, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkut, G.; Madueno, E.; Maltournam, A.; Ma

ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,

A:Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AHL469

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-361 <GLA>

A:Cross-references: GB:AL592022; PIDN:G16412724; GSPDB:GN00178

A:Experimental source: strain Clij11262

C:Genetics:

A:Gene: lin0295

Query Match 47.3%; Score 44; DB 2; Length 361;
Best Local Similarity 42.9%; Pred. No. 15;

Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 IDIKQFWYTPPT 15
DB 310 VTIHDVWVTPGT 323

RESULT 9

T27241

hypothetical protein Y57G11C.31 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27241

R:McMurray, A.
Submitted to the EMBL Data Library, September 1997

A:Reference number: Z20330

A:Accession: T27241

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-425 <WII>

A:Cross-references: EMBL:Z99281; PIDN:CAH16530.1; GSPDB:GN00022; CESP:Y57G11C.31

A:Experimental source: clone Y57G11C

C:Genetics:

A:Gene: CESP:Y57G11C.31

A:Map position: 4

A:introns: 160/3; 187/2; 309/2; 361/3

Query Match 47.3%; Score 44; DB 2; Length 425;
Best Local Similarity 46.7%; Pred. No. 18;

Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 WIDIKQFWYTPPT 15
DB 146 WPEIQDFMTPPSLY 160

RESULT 10

A89958

acetyl-CoA synthetase [imported] - *Staphylococcus aureus* (strain N315)

C:Species: *Staphylococcus aureus*

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 17-May-2002

C:Accession: A89958

R:Kuroda, M.; Ohta, T.; Uchiyama, T.; Baba, T.; Yuzawa, H.; Kobayashi, T.; Cui, L.; Ogura, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.,
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayaishi, H.; Hiratake, K.
Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: A89958

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-568 <KUR>

A:Cross-references: GB:BA000018; PID:G13701528; PIDN:BAH42822.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: acsA

C:Superfamily: acetate-CoA ligase; acetate-CoA ligase homology

Query Match 47.3%; Score 44; DB 2; Length 568;
Best Local Similarity 58.3%; Pred. No. 24;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 DIKQFWYTPPT 14
DB 292 DFKVTIWTAPT 303

RESULT 11

S69206

regulator protein white collar 1 - *Neurospora crassa*

C:Species: *Neurospora crassa*

C:Date: 21-Apr-1997 #sequence_revision 09-May-1997 #text_change 11-Jan-2002

C:Accession: S69206

R:Ballarín, P.; Vitorioso, P.; Magrelli, A.; Talora, C.; Cabibbo, A.; Macino, G.

EMBO J. 15, 1650-1657, 1996

A:Title: White collar-1, a central regulator of blue light responses in *Neurospora*, is

A:Reference number: S69206; MUID:96203083; PMID:8612589

A:Accession: S69206

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1154 <BAL>

A:Cross-references: EMBL:X94300; NID:G1279576; PID:G1480115

C:Genetics:

A:introns: 967/3

C:Superfamily: GATA-type zinc finger homology

C:Keywords: zinc finger

F:932-991/Domain: GATA-type zinc finger homology <GZF>

Query Match 46.8%; Score 43.5; DB 2; Length 1154;
Best Local Similarity 66.7%; Pred. No. 65;

Matches 8; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 3 DIKQFWYTPPT 14
DB 529 DIGQYIW-TPPT 539

RESULT 12

T14669

P-100 protein - *Yersinia pestis* plasmid pMT1

C:Species: *Yersinia pestis*

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T14669

R:Hu, P.; Elliott, J.; McCreedy, P.; Skowronski, E.; Garnea, J.; Kobayashi, A.; Carrano

submitted to the EMBL Data Library, March 1998

A:Description: Structural organization of virulence determinants in three *Yersinia pestis*

A:Reference number: Z18168

A:Accession: T14669

A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-402 <HUP>
 A:Cross-References: EMBL:AF053947; NID:G2996286; PID:G2996304; PIDN:AAC13184.1
 C:Genetics:
 A:Genome: plasmid PMT1

Query Match 46.2%; Score 43; DB 2; Length 402;
 Best Local Similarity 63.6%; Pred. No. 24;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 KQFWYTPPTF 15
 |||||
 Db 38 KORVWVVAPTY 48

RESULT 13

hypothetical protein Y1030 - Yersinia pestis plasmid PMT1

C:Species: Yersinia pestis
 C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 22-Oct-1999

C:Accession: T14946
 C:Residues: 1-418 <LIN>

A:Title: Complete DNA sequence and detailed analysis of the Yersinia pestis KIMS plasmid

A:Reference number: Z18268; MUID:99043898; PMID:9826348

A:Accession: T14946
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-418 <LIN>

A:Cross-References: EMBL:AF074611; NID:G3883003; PID:G3883031; PIDN:AAC82691.1

C:Genetics:

A:Gene: Y1030

A:Genome: plasmid PMT1

Query Match 46.2%; Score 43; DB 2; Length 418;
 Best Local Similarity 63.6%; Pred. No. 25;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 KQFWYTPPTF 15
 |||||
 Db 54 KORVWVVAPTY 64

RESULT 14

gag polyprotein - human immunodeficiency virus type 1 (strain Ant70)

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 07-May-1999

C:Accession: A53034

R:Vanden Haesevelde, M.; Decourt, J.L.; De Leys, R.J.; Vanderborcht, B.; van der Groen, U. Virol. 68, 1586-1596, 1994

A:Title: Genomic cloning and complete sequence analysis of a highly divergent African hu

A:Reference number: A53034; MUID:94149849; PMID:8107220

A:Accession: A53034

A:Status: preliminary

A:Molecule type: genomic RNA

A:Residues: 1-863 <VAN>

A:Cross-References: GB:L02587

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: polyprotein

Query Match 46.2%; Score 43; DB 2; Length 863;
 Best Local Similarity 50.0%; Pred. No. 57;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 WIDIKQVWY 10
 |||||
 Db 672 WIDITKMLWY 681

RESULT 15

S49197

envelope protein precursor - human immunodeficiency virus type 1 (fragment)
 C:Species: human immunodeficiency virus type 1, HIV-1
 C>Date: 19-Mar-1997 #sequence_revision 24-Jul-1997 #text_change 26-Aug-1999

C:Accession: S49197

R:Charneau, P.; Borman, A.M.; Quillent, C.; Guetard, D.; Charette, S.; Cohen, J.; Remy,

submitted to the EMBL Data Library, July 1994

A:Description: Isolation and envelope sequence of a highly divergent HIV-1 isolate; defi

A:Reference number: S49197

A:Accession: S49197

A:Molecule type: DNA

A:Residues: 1-877 <CHA>

A:Cross-References: EMBL:X80020; NID:G510516; PIDN:CA56323.1; PID:G510517

A:Experimental source: isolate VAV

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: glycoprotein; capsid protein; coat protein; transmembrane protein

F:1-30/Domain: signal sequence #status predicted <Sig>

F:31-535/Product: coat protein gp120 #status predicted <CP1>

F:536-877/Product: coat protein gp41 #status predicted <CP2>

F:698-716/Domain: transmembrane #status predicted <TMN>

F:59,88,139,148,159,184,188,198,230,235,242,263,270,277,292,302,333,345,357,367,396,404,

Query Match 46.2%; Score 43; DB 2; Length 877;
 Best Local Similarity 50.0%; Pred. No. 58;

Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 WIDIKQVWY 10
 |||||
 Db 686 WIDITKMLWY 695

Search completed: July 12, 2004, 14:11:20
 Job time : 8.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:10:33 ; Search time 26.5116 Seconds

(without alignments)
176,480 Million cell updates/sec

Title: US-10-010-667a-2_COPY_277_291

Sequence: 93
1 WIDIKQFWYTPPTF 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1279676 seqs, 311918243 residues

Total number of hits satisfying chosen parameters: 1279676

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	100.0	15	14	US-10-011-095-21
2	93	100.0	15	14	US-10-010-667a-21
3	93	100.0	15	14	US-10-165-044-39
4	93	100.0	267	9	US-09-747-835A-50
5	93	100.0	267	9	US-09-747-835A-51
6	93	100.0	267	12	US-10-312-312-50
7	93	100.0	267	12	US-10-312-312-51
8	93	100.0	339	9	US-09-759-143-879
9	93	100.0	339	9	US-09-780-669-879
10	93	100.0	339	9	US-09-822-827-879
11	93	100.0	339	9	US-09-802-520-11
12	93	100.0	339	9	US-09-895-793-879
13	93	100.0	339	9	US-09-895-814-879
14	93	100.0	339	12	US-10-205-267-13
15	93	100.0	339	12	US-10-408-009-2

16	93	100.0	339	13	US-10-012-896-879	Sequence 879, App
17	93	100.0	339	14	US-10-011-095-21	Sequence 2, Appli
18	93	100.0	339	14	US-10-010-667a-2	Sequence 2, Appli
19	93	100.0	339	14	US-10-205-823-397	Sequence 397, App
20	93	100.0	339	14	US-10-144-678A-879	Sequence 879, App
21	93	100.0	339	14	US-10-294-025-879	Sequence 879, App
22	93	100.0	339	15	US-10-239-607-37	Sequence 37, Appl
23	93	100.0	339	15	US-10-235-027-714	Sequence 714, App
24	93	100.0	339	15	US-10-295-027-1347	Sequence 1347, Ap
25	93	100.0	368	12	US-10-425-114-72779	Sequence 72779, A
26	93	100.0	375	14	US-10-165-044-2	Sequence 2, Appli
27	45	48.4	268	9	US-09-854-816-39	Sequence 39, Appl
28	45	48.4	269	9	US-09-854-816-33	Sequence 33, Appl
29	45	48.4	883	12	US-10-282-122A-57398	Sequence 57398, A
30	45	48.4	854	14	US-10-369-294-17	Sequence 17, Appl
31	45	48.4	871	15	US-10-346-000A-9	Sequence 9, Appli
32	44.5	47.8	553	12	US-10-282-122A-77905	Sequence 77905, A
33	44.5	47.8	662	15	US-10-369-493-921	Sequence 921, App
34	44	47.3	70	16	US-10-437-963-139012	Sequence 139012, A
35	44	47.3	91	12	US-10-424-599-232688	Sequence 232688, A
36	44	47.3	103	16	US-10-437-963-115402	Sequence 115402, A
37	44	47.3	403	14	US-10-301-997-83	Sequence 83, Appl
38	44	47.3	568	9	US-09-815-242-5492	Sequence 5492, Ap
39	44	47.3	568	9	US-09-815-242-12629	Sequence 12629, A
40	44	47.3	568	12	US-10-282-122A-44222	Sequence 44222, A
41	43	46.2	71	16	US-10-437-963-115402	Sequence 115402, A
42	43	46.2	89	14	US-10-058-053A-62	Sequence 62, Appl
43	43	46.2	98	12	US-10-424-599-193800	Sequence 193800, A
44	43	46.2	119	12	US-10-257-174-41	Sequence 41, Appl
45	43	46.2	147	9	US-09-873-134-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-10-011-095-21
Sequence 21, Application US/10011095
Publication No. US20030045682A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Hubert, Rene S.
APPLICANT: Leong, Kahan
APPLICANT: Raitano, Arthur B.
APPLICANT: Saltman, Douglas C.
APPLICANT: Mitchell, Steve Chappel
TITLE OF INVENTION: ANTIBODIES IMMUNOSPECIFIC FOR STEAP1 (AS AMENDED)
FILE REFERENCE: 511582001610
CURRENT APPLICATION NUMBER: US/10/011,095
PRIOR FILING DATE: 2001-12-06
PRIOR APPLICATION NUMBER: 09/323,873
PRIOR FILING DATE: 1999-06-01
PRIOR APPLICATION NUMBER: 60/087,520
PRIOR FILING DATE: 1998-06-01
PRIOR APPLICATION NUMBER: 60/091,183
PRIOR FILING DATE: 1998-06-30
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21
LENGTH: 15
TYPE: PPT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: STEAP-1 PEPTIDE
US-10-011-095-21
Query Match 100.0%; Score 93; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 WIDIKQFWYTPPTF 15
Db 1 WIDIKQFWYTPPTF 15

```

RESULT 2
US-10-010-667A-21
; Sequence 21, Application US/10010667A
; Publication No. US20030055217A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Hubert, Rene S.
; APPLICANT: Leong, Kahan
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Saffran, Douglas C.
; APPLICANT: Mitchell, Steve Chappell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 511582001601
; CURRENT APPLICATION NUMBER: US/10/010,667A
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: STEAP-1 PEPTIDE
US-10-010-667A-21

Query Match          100.0%; Score 93; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches    15; Conservative    0; Mismatches    0; Indels    0; Gaps    0

QY      1 WIDIKQFYWYTPTF 15
        |||
Db       1 WIDIKQFYWYTPTF 15

RESULT 3
US-10-165-044--39
; Sequence 39, Application US/10165044
; Publication No. US20030149531A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas Saffran
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Steven Chappel Mitchell
; APPLICANT: Mary Faris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 51158-20016.02
; CURRENT APPLICATION NUMBER: US/10/165,044
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: US 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: US 60/091,183
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 09/323,873
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 09/455,486
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: WO 99/62941
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/33040
; PRIOR FILING DATE: 2000-12-06

```

```

; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-165-044-39

Query Match          100.0%; Score 93; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDIKQFVWYTPPTF 15
      |||||
Db      1 WIDIKQFVWYTPPTF 15

RESULT 4
US-09-747-835A-50
; Sequence 50, Application US/09747835A
; Patent No. US20020146692A1
; GENERAL INFORMATION:
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Dmarac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
; FILE REFERENCE: HYS-37C1P
; CURRENT APPLICATION NUMBER: US/09/747,835A
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US 09/729,739
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US 09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-747-835A-50

Query Match          100.0%; Score 93; DB 9; Length 267;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDIKQFVWYTPPTF 15
      |||||
Db      211 WIDIKQFVWYTPPTF 225

RESULT 5
US-09-747-835A-51
; Sequence 51, Application US/09747835A
; Patent No. US20020146692A1
; GENERAL INFORMATION:
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua

```



```

; APPLICANT: Zhou, Ping
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Dimanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
; FILE REFERENCE: HXS-37CIP
; CURRENT APPLICATION NUMBER: US/09/747,835A
; CURRENT FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US 09/729,739
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US 09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 51
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-747-835A-51

Query Match          100.0%; Score 93; DB 9; Length 267;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDIKQFWYTPPTF 15
        |||||
Db      211 WIDIKQFWYTPPTF 225

RESULT 6
US-10-312-312-50
; Sequence 50, Application US/10312312
; Publication No. US20040068097A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Dimanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
; FILE REFERENCE: 21272-015-061/HXS-37CIP
; CURRENT APPLICATION NUMBER: US/10/312,312
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 09/729,739
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US 09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 51
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-312-312-51
```

```

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-312-312-50

Query Match          100.0%; Score 93; DB 12; Length 267;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDIKQFWYTPPTF 15
        |||||
Db      211 WIDIKQFWYTPPTF 225

RESULT 7
US-10-312-312-51
; Sequence 51, Application US/10312312
; Publication No. US20040068097A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Dimanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
; FILE REFERENCE: 21272-015-061/HXS-37CIP
; CURRENT APPLICATION NUMBER: US/10/312,312
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 09/729,739
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US 09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 51
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-312-312-51

Query Match          100.0%; Score 93; DB 12; Length 267;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDIKQFWYTPPTF 15
        |||||
Db      211 WIDIKQFWYTPPTF 225

RESULT 8
US-09-759-143-879
; Sequence 879, Application US/09759143
; Patent No. US20020022248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
```

```

; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-143-879
```

```

Query Match      100.0%; Score 93; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 WIDIKQFWYTPPTF 15
        |||||
Db      277 WIDIKQFWYTPPTF 291
```

```

RESULT 9
US-09-780-669-879
; Sequence 879, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780,669
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-780-669-879
```

```

Query Match      100.0%; Score 93; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 WIDIKQFWYTPPTF 15
        |||||
Db      277 WIDIKQFWYTPPTF 291
```

```

RESULT 10
US-09-822-827-879
; Sequence 879, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-822-827-879
```

```

Query Match      100.0%; Score 93; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 WIDIKQFWYTPPTF 15
        |||||
Db      277 WIDIKQFWYTPPTF 291
```

```

RESULT 11
US-09-802-520-11
; Sequence 11, Application US/09802520
; Publication No. US20020187472A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Paris, Mary
; APPLICANT: Chen, Hui-Mei
; APPLICANT: Ison, Craig H.
; TITLE OF INVENTION: STEAP-RELATED PROTEIN
; FILE REFERENCE: PC-0037 US
; CURRENT APPLICATION NUMBER: US/09/802,520
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PERL Program
; SEQ ID NO: 11
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020187472A1 96572948
US-09-802-520-11
```

```

Query Match      100.0%; Score 93; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 WIDIKQFWYTPPTF 15
        |||||
Db      277 WIDIKQFWYTPPTF 291
```

```

RESULT 12
US-09-895-793-879
```

```
; Sequence 879, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiaqichun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiaqiang, Yugu
; APPLICANT: Katos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hurai, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-895-793-879

Query Match      100.0%; Score 93; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDIKQFWMYTPPTF 15
DB      277 WIDIKQFWMYTPPTF 291

RESULT 13
US-09-895-814-879
; Sequence 879, Application US/09895814
; Publication No. US20020193236A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiaqichun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiaqiang, Yugu
; APPLICANT: Katos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hurai, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
```

```
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-895-814-879

Query Match      100.0%; Score 93; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDIKQFWMYTPPTF 15
DB      277 WIDIKQFWMYTPPTF 291

RESULT 14
US-10-205-267-13
; Sequence 13, Application US/10205267
; Publication No. US20030064397A1
; GENERAL INFORMATION:
; APPLICANT: Spence, Kimberly M.
; APPLICANT: Rickett, Paula K.
; APPLICANT: Lal, Preeti G.
; APPLICANT: Ison, Craig H.
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN PROSTATE AND
; TITLE OF INVENTION: TUMORS
; FILE REFERENCE: PV-0008 CIP
; CURRENT APPLICATION NUMBER: US/10/205,267
; CURRENT FILING DATE: 2002-07-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 339
; TYPE: PRT
; ORGANISM:
; PUBLICATON INFORMATION:
; DATABASE ACCESSION NUMBER: Genbank ID No: 96572948
; US-10-205-267-13

Query Match      100.0%; Score 93; DB 12; Length 339;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDIKQFWMYTPPTF 15
DB      277 WIDIKQFWMYTPPTF 291

RESULT 15
US-10-408-009-2
; Sequence 2, Application US/10408009
; Publication No. US20040072196A1
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Stephen C. Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 511582001603
; CURRENT APPLICATION NUMBER: US/10/408,009
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 09/455,486
; PRIOR FILING DATE: 1999-12-06
```

```
; PRIOR APPLICATION NUMBER: 09/323,873
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ. ID NOS: 36
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ. ID NO. 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-009-2
```

```
Query Match      100.0%; Score 93; DB 12; Length 339;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 WIDIKQFWYTPPTF 15
        |||||
Db      277 WIDIKQFWYTPPTF 291
```

Search completed: July 12, 2004, 14:27:49
Job time : 26.5116 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 14:04:08 ; Search time 9.24419 Seconds

(without alignments)
83.770 Million cell updates/sec

Title: US-10-010-667a-2_COPY_277_291
Sequence: 93
1 WIDIKQFWYTPPTF 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents AA:*
2: /cgn2_6/prodata/2/1aa/5a COMB.pep:*
3: /cgn2_6/prodata/2/1aa/5b COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6a COMB.pep:*
5: /cgn2_6/prodata/2/1aa/6b COMB.pep:*
6: /cgn2_6/prodata/2/1aa/PCTUS COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	100.0	15	4	US-09-323-873A-21 Sequence 21, Appl
2	93	100.0	339	4	US-09-323-873A-2 Sequence 2, Appl
3	93	100.0	339	4	US-09-685-166A-879 Sequence 879, App
4	45	48.4	268	3	US-08-965-056-39 Sequence 39, Appl
5	45	48.4	269	3	US-08-965-056-33 Sequence 33, Appl
6	45	48.4	353	1	US-08-118-270-45 Sequence 45, Appl
7	45	48.4	353	5	PCT-US93-08528-45 Sequence 45, Appl
8	45	48.4	489	4	US-09-134-000C-4888 Sequence 4888, Ap
9	45	48.4	854	4	US-09-206-551-17 Sequence 17, Appl
10	45	48.4	865	3	US-07-956-483-13 Sequence 13, Appl
11	45	48.4	887	3	US-08-472-240A-4 Sequence 4, Appl
12	44	47.3	403	4	US-08-887-534A-83 Sequence 83, Appl
13	44	47.3	403	4	US-09-527-431-83 Sequence 83, Appl
14	43	46.2	351	4	US-08-817-441-47 Sequence 47, Appl
15	43	46.2	862	2	US-09-206-551-15 Sequence 15, Appl
16	43	46.2	873	2	US-08-912-129A-61 Sequence 61, Appl
17	43	46.2	877	4	US-08-817-441-102 Sequence 102, App
18	42	45.2	268	3	US-08-965-056-26 Sequence 26, Appl
19	42	45.2	268	3	US-08-965-056-68 Sequence 68, Appl
20	42	45.2	268	3	US-08-965-056-71 Sequence 71, Appl
21	42	45.2	446	4	US-09-543-681A-5890 Sequence 5890, Ap
22	42	45.2	533	4	US-09-107-532A-7241 Sequence 7241, Ap
23	41	44.1	231	3	US-08-448-489-19 Sequence 19, Appl
24	41	44.1	237	3	US-08-388-353-642 Sequence 642, App
25	41	44.1	237	3	US-08-488-551B-642 Sequence 642, App
26	41	44.1	243	3	US-08-965-056-73 Sequence 73, Appl
27	41	44.1	268	3	US-08-965-056-8 Sequence 8, Appl

28	41	44.1	268	3	US-08-965-056-13 Sequence 13, Appl
29	41	44.1	268	3	US-08-965-056-14 Sequence 14, Appl
30	41	44.1	268	3	US-08-965-056-70 Sequence 70, Appl
31	41	44.1	268	3	US-08-965-056-72 Sequence 72, Appl
32	41	44.1	269	3	US-08-965-056-22 Sequence 22, Appl
33	41	44.1	269	3	US-08-965-056-25 Sequence 25, Appl
34	41	44.1	269	3	US-08-965-056-37 Sequence 37, Appl
35	41	44.1	270	3	US-08-965-056-67 Sequence 67, Appl
36	41	44.1	351	1	US-08-470-202-46 Sequence 46, Appl
37	41	44.1	351	1	US-08-471-770-46 Sequence 46, Appl
38	41	44.1	351	2	US-08-468-059-46 Sequence 46, Appl
39	41	44.1	351	3	US-09-109-916-46 Sequence 46, Appl
40	41	44.1	351	4	US-09-886-156-46 Sequence 46, Appl
41	41	44.1	351	4	US-09-886-149-46 Sequence 46, Appl
42	41	44.1	351	4	US-09-886-150-46 Sequence 46, Appl
43	41	44.1	351	4	US-09-886-159-46 Sequence 46, Appl
44	41	44.1	602	3	US-09-257-490-15 Sequence 15, Appl
45	41	44.1	602	3	US-09-257-490-15 Sequence 15, Appl

ALIGNMENTS

```
RESULT 1
US-09-323-873A-21
; Sequence 21, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 129.16TSU2
; CURRENT APPLICATION NUMBER: US/09/323,873A
; CURRENT FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: STEAP-1 PEPTIDE
US-09-323-873A-21
Query Match      100.0%; Score 93; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.5e+08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 WIDIKQFWYTPPTF 15
Db      1 WIDIKQFWYTPPTF 15
RESULT 2
US-09-323-873A-2
; Sequence 2, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Rene S. Hubert
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell
```

```
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 129.160SU2
; CURRENT APPLICATION NUMBER: US/09/323,873A
; CURRENT FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-09-323-873A-2

Query Match      100.0%; Score 93; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 8.3e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDIKQFVWYTPPTF 15
Db      277 WIDIKQFVWYTPPTF 291

RESULT 3
US-09-685-166A-879
; Sequence 879, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jianshun
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jjiang, Yugu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-685-166A-879

Query Match      100.0%; Score 93; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 8.3e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDIKQFVWYTPPTF 15
Db      277 WIDIKQFVWYTPPTF 291

RESULT 4
US-08-965-056-39
; Sequence 39, Application US/08965056

Patent No. 6271198
; GENERAL INFORMATION:
; APPLICANT: Andrew C. Braisted
; APPLICANT: J. Kevin Judice
; APPLICANT: Robert S. McDowell
; APPLICANT: J. Christopher Phelan
; APPLICANT: Melissa A. Starovasnik
; APPLICANT: James A. Wells
; TITLE OF INVENTION: Constrained Helical Peptides and Methods of
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/965,056
; FILING DATE: 05-NO. 6271198-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Ph.D., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P1005R2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 268 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-965-056-39

Query Match      48.4%; Score 45; DB 3; Length 268;
Best Local Similarity 60.0%; Pred. No. 17;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 WIDIKQFWY 10
Db      202 WFDITQWLWY 211

RESULT 5
US-08-965-056-33
; Sequence 33, Application US/08965056
; Patent No. 6271198
; GENERAL INFORMATION:
; APPLICANT: Andrew C. Braisted
; APPLICANT: J. Kevin Judice
; APPLICANT: Robert S. McDowell
; APPLICANT: J. Christopher Phelan
; APPLICANT: Melissa A. Starovasnik
; APPLICANT: James A. Wells
; TITLE OF INVENTION: Constrained Helical Peptides and Methods of
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965 056
FILING DATE: 05-NO. 6271198-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, PhD., Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: P1005R2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-8674
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 269 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-965-056-33

Query Match 48.4%; Score 45; DB 3; Length 269;
Best Local Similarity 60.0%; Pred. No. 17;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WIDIKQFWY 10
DB 203 WFDITQWLVY 212

RESULT 6
US-08-118-270-45
Sequence 45, Application US/08118270
Patent No. 5508384
GENERAL INFORMATION:
APPLICANT: Murphy, Randall B.
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSER: BROWDY AND NEWMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/118,270
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY=2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 353 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide

US-08-118-270-45

Query Match 48.4%; Score 45; DB 1; Length 353;
Best Local Similarity 33.3%; Pred. No. 22;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 WIDIKQFWYPTPTF 15
DB 56 WVELNFTWHPWAF 70

RESULT 7
PCT-US93-08528-45
Sequence 45, Application PC/TUS9308528
GENERAL INFORMATION:
APPLICANT: New York University
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSER: BROWDY AND NEWMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08528
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY=2 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 353 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-08528-45

Query Match 48.4%; Score 45; DB 5; Length 353;
Best Local Similarity 33.3%; Pred. No. 22;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 WIDIKQFWYPTPTF 15
DB 56 WVELNFTWHPWAF 70

RESULT 8
US-09-134-000C-4888
Sequence 4888, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucelte-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
FILE REFERENCE: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
CURRENT APPLICATION NUMBER: US/09/134,000C

```

; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4888
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-4888

Query Match
Best Local Similarity 48.4%; Score 45; DB 4; Length 489;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MIDIKQFVWYTP 13
Db 450 WPEITQIWLAP 462

RESULT 9
US-09-206-551-17
; Sequence 17, Application US/09206551B
; Patent No. 6521739
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice H.
; APPLICANT: Gao, Feng
; APPLICANT: Marx, Preston A.
; APPLICANT: Shaw, George M.
; APPLICANT: Smith, Stephen M.
; APPLICANT: Georges-Courbot, Marie Claude
; APPLICANT: Lu, Chang Yong
; TITLE OF INVENTION: Complete Genome Sequences of a Simian
; TITLE OF INVENTION: Immunodeficiency Virus from a Red-Capped
; FILE REFERENCE: D6286
; CURRENT APPLICATION NUMBER: US/09/206,551B
; CURRENT FILING DATE: 1998-12-07
; NUMBER OF SEQ ID NOS: 58
; SEQ ID NO 17
; LENGTH: 854
; TYPE: PRT
; ORGANISM: Simian immunodeficiency virus
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of homologous region of
; US-09-206-551-17

Query Match
Best Local Similarity 48.4%; Score 45; DB 4; Length 854;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIDIKQFVWY 10
Db 670 WPDITQWLWY 679

RESULT 10
US-07-956-483-13
; Sequence 13, Application US/07956483
; Patent No. 6261799
; GENERAL INFORMATION:
; APPLICANT: KIENY, Marie-Paule
; TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/956,483
; FILING DATE: 31-DEC-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 92/19742
; FILING DATE: 12-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 05392
; FILING DATE: 02-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feuty, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 017753-005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 865 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-07-956-483-13
```

```

Query Match
Best Local Similarity 48.4%; Score 45; DB 3; Length 865;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 1 MIDIKQFVWY 10
Db 681 WPDITQWLWY 690

RESULT 11
US-08-472-240A-4
; Sequence 4, Application US/08472240A
; Patent No. 6284248
; GENERAL INFORMATION:
; APPLICANT: KIENY, Marie-Paule
; TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,240A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/956,483
; FILING DATE: 31-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-055
; TELECOMMUNICATION INFORMATION:
```


TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 887 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Protein
LOCATION: 1..854
US-08-472-240A-4

Query Match 48.4%; Score 45; DB 3; Length 887;
Best Local Similarity 60.0%; Pred. No. 56;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIDIKOFWY 10
DB 703 MFDITQWLY 712

RESULT 12
US-08-887-534A-83
Sequence 83, Application US/08887534A
Patent No. 6455323
GENERAL INFORMATION:
APPLICANT: Holden, David W.
TITLE OF INVENTION: ANTI-BACTERIAL METHODS AND MATERIALS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Garstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/887,534A
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 28341/33996
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-6300
TELEX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 403 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-887-534A-83

Query Match 47.3%; Score 44; DB 4; Length 403;
Best Local Similarity 58.3%; Pred. No. 36;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 DIKOFWYTPPT 14
DB 305 DFKYTIWYAPT 316

RESULT 13

US-09-527-431-83
Sequence 83, Application US/09527431
Patent No. 6485899
GENERAL INFORMATION:
APPLICANT: Holden, David W.
TITLE OF INVENTION: ANTI-BACTERIAL METHODS AND MATERIALS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Garstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/527,431
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/887,534
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 28341/33996
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: (312) 474-6600
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 403 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-527-431-83

Query Match 47.3%; Score 44; DB 4; Length 403;
Best Local Similarity 58.3%; Pred. No. 36;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 DIKOFWYTPPT 14
DB 305 DFKYTIWYAPT 316

RESULT 14
US-08-817-441-47
Sequence 47, Application US/08817441
Patent No. 6399294
GENERAL INFORMATION:
APPLICANT: CHARNEAU, PIERRE
APPLICANT: CLAVEL, FRANCOISE
APPLICANT: BORMAN, ANDREW
APPLICANT: OUILLENT, CAROLINE
APPLICANT: GUETARD, DENISE
APPLICANT: MONTAGNER, LUC
APPLICANT: DONTON DE SAINT-MARTIN, JACQUELINE
APPLICANT: COHEN, JAOUES
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF HIV-1 TYPE (OR
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner, L.L.P.
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: DC

COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,441
FILING DATE: 11-JUL-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 95/01391
FILING DATE: 20-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9412554
FILING DATE: 20-OCT-1994
APPLICATION NUMBER: FR 9502526
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03260.6005-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 351 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-817-441-47

Query Match 46.2%; Score 43; DB 4; Length 351;
Best Local Similarity 50.0%; Pred. No. 45;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 WIDIKQFWY 10
|:|:|:|:
Db 160 WIDITKMLWY 169

RESULT 15
US-09-206-551-15
; Sequence 15, Application US/09206551B
; Patent No. 6521739
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice H.
; APPLICANT: Gao, Feng
; APPLICANT: Marx, Preston A.
; APPLICANT: Shaw, George M.
; APPLICANT: Smith, Stephen M.
; APPLICANT: Georges-Courbot, Marie Claude
; APPLICANT: Lu, Chang Yong
; TITLE OF INVENTION: Complete Genome Sequences of a Simian
; TITLE OF INVENTION: Immunodeficiency Virus from a Red-Capped
; FILE REFERENCE: D6286
; CURRENT APPLICATION NUMBER: US/09/206,551B
; CURRENT FILING DATE: 1998-12-07
; NUMBER OF SEQ ID NOS: 58
; SEQ ID NO 15
; LENGTH: 862
; TYPE: PRT
; ORGANISM: Simian immunodeficiency virus
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of homologous region of
; OTHER INFORMATION: O_ANT10 lentiviral env protein
US-09-206-551-15

Query Match

46.2%; Score 43; DB 4; Length 862;

Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 WIDIKQFWY 10
|:|:|:|:
Db 671 WIDITKMLWY 680

Search completed: July 12, 2004, 14:12:24
JOB time : 9.24419 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2004, 13:51:07 ; Search time 32.9651 Seconds
(without alignments)
128.567 Million cell updates/sec

Title: US-10-010-667a-2_COPY_277_291
Perfect score: 93
Sequence: 1 WIDIKQFWYTPPTF 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: geneseqp1980s:*
- 2: geneseqp1900s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003s:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	93	100.0	15	3	AAy58200 Human STR
2	93	100.0	15	4	AAE02788 Extracell
3	93	100.0	267	6	ABU60886 Human G P
4	93	100.0	267	6	ABU60887 Human G P
5	93	100.0	339	3	AAy58194 Human STR
6	93	100.0	339	4	AAm01282 P789P ami
7	93	100.0	339	4	AAu69927 Human pro
8	93	100.0	339	4	AAm78845 Human pro
9	93	100.0	339	4	AAu71818 Human pro
10	93	100.0	339	5	ABB95387 Human P78
11	93	100.0	339	5	ABG61813 Human P78
12	93	100.0	339	6	ABU98383 Prostate
13	93	100.0	339	6	ABU98414 STEAP-1 v
14	93	100.0	339	6	ABU98430 STEAP-1 v
15	93	100.0	339	6	ABU98430 STEAP-1 v
16	93	100.0	339	7	ABU63313 Prostate
17	93	100.0	339	7	ABD75573 Prostate
18	93	100.0	339	7	ABD75573 Prostate
19	93	100.0	339	7	ABD75573 Prostate
20	93	100.0	339	7	ABD75573 Prostate
21	93	100.0	339	7	ABD75573 Prostate
22	93	100.0	339	7	ABD75573 Prostate
23	93	100.0	339	7	ABD75573 Prostate
24	93	100.0	339	7	ABD75573 Prostate
25	93	100.0	339	7	ABD75573 Prostate

26	45	48.4	269	2	AAy22837 SEQ ID NO
27	45	48.4	269	5	ABG68308 Envelope
28	45	48.4	269	6	ABU57715 Human Imm
29	45	48.4	353	2	AAu48724 G-Protein
30	45	48.4	353	2	AAu48724 G-Protein
31	45	48.4	483	6	ABU29474 Protein e
32	45	48.4	846	3	AAu69345 HIV-1 non
33	45	48.4	861	2	AAu69345 HIV-1 gpl
34	45	48.4	861	2	AAu69345 HIV-1 gpl
35	45	48.4	861	2	AAu69345 HIV-1 gpl
36	45	48.4	861	2	AAu69345 HIV-1 gpl
37	45	48.4	861	2	AAu69345 HIV-1 gpl
38	45	48.4	861	2	AAu69345 HIV-1 gpl
39	45	48.4	861	2	AAu69345 HIV-1 gpl
40	45	48.4	861	2	AAu69345 HIV-1 gpl
41	45	48.4	861	2	AAu69345 HIV-1 gpl
42	45	48.4	861	2	AAu69345 HIV-1 gpl
43	45	48.4	861	2	AAu69345 HIV-1 gpl
44	45	48.4	861	2	AAu69345 HIV-1 gpl
45	45	48.4	861	2	AAu69345 HIV-1 gpl

ALIGNMENTS

RESULT 1	AAy58200	standard; peptide; 15 AA.
ID	AAy58200	standard; peptide; 15 AA.
XX	AAy58200	standard; peptide; 15 AA.
AC	AAy58200	standard; peptide; 15 AA.
XX	AAy58200	standard; peptide; 15 AA.
DT	14-MAR-2000	(first entry)
XX	14-MAR-2000	(first entry)
DE	Human STRAP-1 peptide, corresponding to STRAP-1 extracellular region 3.	
XX	Human STRAP-1 peptide, corresponding to STRAP-1 extracellular region 3.	
KW	Serpinin transmembrane antigen of the prostate; STRAP-1; prostate;	
KW	transmembrane domain; type IIIa membrane protein; expression; cancer;	
KW	prostate cancer; bladder cancer; colon cancer; pancreatic cancer;	
KW	ovarian cancer; tumour antigen; immunisation; immune response; cellular;	
KW	humoral; anticancer vaccine; antibody; detection; diagnosis; prognosis;	
KW	monitoring; susceptibility; therapeutic inhibitor; drug targeting;	
XX	recombinant protein.	
OS	Synthetic.	
OS	Homo sapiens.	
XX	Homo sapiens.	
PN	WO962941-A2.	
XX	WO962941-A2.	
PD	09-DEC-1999.	
XX	09-DEC-1999.	
PF	01-JUN-1999;	99WO-US012157.
XX	01-JUN-1999;	99WO-US012157.
PR	01-JUN-1998;	98US-0087520P.
PR	30-JUN-1998;	98US-0091183P.
XX	30-JUN-1998;	98US-0091183P.
PA	(UROC-) UROGENESYS INC.	
PA	(AFAR/) AFAR D E.	
PA	(HUBER/) HUBERT R S.	
PA	(LEON/) LEON K.	
PA	(RAIT/) RAITANO A B.	
PA	(SAFE/) SAFRAN D C.	
PI	AFAR DE, Hubert RS, Leong K, Raitano AB, Safran DC;	
DR	WPI; 2000-072832/06.	
XX	WPI; 2000-072832/06.	
PT	Novel proteins useful as diagnostic markers and therapeutic targets,	
PT	particularly for prostatic cancer.	
XX	Particularly for prostatic cancer.	
XX	Disclosure; Page 22; 83pp; English.	
PS	Sequences AAy58198-Y58200 represent synthetic peptides that correspond to	
XX	the extracellular regions of STRAP-1 (serpinin transmembrane antigen of	
CC	the extracellular regions of STRAP-1 (serpinin transmembrane antigen of	

the prostate, AAV58194). These peptides were used to raise monoclonal anti-STRAP-1 antibodies. STRAP-1 is the prototype member of the STRAP family of proteins (AAV58194-Y58197) which exhibit a high degree of structural conservation, but which show no significant structural homology to known human proteins. The STRAP-1 gene has been localised to chromosome 7p22. STRAP-1 is thought to be a type IIIa membrane protein and is expressed predominantly in prostate cells in normal human tissues. Structurally, STRAP-1 is a 339 amino acid protein characterised by six transmembrane domains and intracellular N- and C-termini, suggesting that it folds in a "serpentine" manner into three extracellular and two intracellular loops. STRAP-1 mRNA and protein expression is maintained at high levels and throughout all stages of prostate cancer. STRAP-1 mRNA and/or protein is also overexpressed in certain other cancers, including bladder, colon, pancreatic and ovarian cancer. The function of the STRAP protein is not known. They may be ion channels (from the presence of six transmembrane domains, a feature which is shared by certain ion channels) or gap-junction proteins (from immunohistochemical staining). STRAP-1 and STRAP-2 are cell-surface tumour antigens. Immunisation with a STRAP protein induces cellular and humoral immune responses against STRAP-expressing cells. STRAP proteins may be used to identify specific-binding agents, to produce anticancer vaccines and to generate specific antibodies. The antibodies may be used for detection, prognosis, and monitoring of cancers (or susceptibility to cancer), as therapeutic inhibitors or to target therapeutic agents to their site of action. STRAP nucleic acids may be used for recombinant protein production, as diagnostic and prognostic reagents, for identifying STRAP-expressing cells for screening inhibitors of STRAP expression and for therapeutic modulation/inhibition of STRAP expression. Since high levels of STRAP proteins are exposed on the cell surface, they are easily targeted by systemically administered agents, and because they are expressed mainly on prostatic epithelial cells, agents targeted to them should have minimal side effects on other tissues

SQ Sequence 15 AA;

Query Match 100.0%; Score 93; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WIDIKQFWWYTPPTF 15
|||||
1 WIDIKQFWWYTPPTF 15

RESULT 2
AAE02788 standard; peptide; 15 AA.

AC AAE02788;

DT 06-AUG-2001 (first entry)

DE Extracellular loop #3 of human STRAP-1, suitable for cloning into pFc.

XX Human, cytosolic; antiproliferative; vaccine; gene therapy;
KW six transmembrane epithelial antigen of the prostate-1; STRAP-1;
KM chromosome 7p22.3; cancer; prostate; colon; bladder; pancreatic; ovarian;
lung; extracellular loop; serpentine transmembrane antigen.

XX Homo sapiens.

XX MO200140276-A2.

XX 07-JUN-2001.

XX 06-DEC-2000; 2000WO-US033040.

XX 06-DEC-1999; 99US-00455486.

XX (UROG-) UROGENESYS INC.

XX Afar DEH, Hubert RS, Raitano AB, Saffran DC, Mitchell SC;
PI Paris M, Jakobovits A;

XX WPI; 2001-367804/38.

XX New STRAP (six transmembrane epithelial antigen of the prostate)
PT proteins, expressed in human cancers, useful for detecting and treating
XX cancer.

XX Example 19; Page 102; 187pp; English.

XX The present invention relates to human six transmembrane epithelial
CC antigen of the prostate (STRAP) protein. STRAP is a member of cell
CC surface serpentine transmembrane antigens. STRAP gene is used in gene
CC therapy, inhibiting the development or progression of a cancer (eg.
CC prostate, colon, bladder, lung, ovarian and pancreatic) expressing STRAP
CC or inhibiting growth or killing cells expressing STRAP in a patient,
CC comprising administering a vaccine composition to the patient. Treating a
CC patient with a cancer that expresses STRAP, or inhibiting growth or
CC killing cells expressing STRAP, comprises administering to the patient a
CC vector encoding single chain monoclonal antibody that comprises the
CC variable domains of the heavy and light chains of the monoclonal antibody
CC that specifically binds to STRAP, such that the vector delivers the
CC single chain monoclonal antibody coding sequence to the cancer cells and
CC the encoded single chain monoclonal antibody is expressed
CC intracellularly. The present sequence is extracellular loop of STRAP-1
CC suitable for cloning into pFc, which is used in the invention. STRAP-1
CC gene is located on chromosome 7p22.3

SQ Sequence 15 AA;

Query Match 100.0%; Score 93; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WIDIKQFWWYTPPTF 15
|||||
1 WIDIKQFWWYTPPTF 15

RESULT 3
ABU60886 standard; protein; 267 AA.

AC ABU60886;

DT 08-MAY-2003 (first entry)

DE Human G protein-coupled receptor-related (GPCR-related) protein #3.

XX Human; G protein-coupled receptor; GPCR; receptor; ophthalmic disease;
KW neurological disease; immunological disease; nephritic disease; cancer;
KW hormonal dysfunction; neoplasm; atherosclerosis; diabetes; chemotherapy;
KW anaemia; irradiation; erythroid precursor; erythroid cell; ulcer; organ;
KW pressure ulcer; vascular insufficiency; surgical wound; traumatic wound;
KW pancreas; liver; intestine; kidney; skin; endothelium; skeletal muscle;
KW smooth muscle; cardiac muscle; vascular tissue.

XX Homo sapiens.

XX US2002146692-A1.

XX 10-OCT-2002.

XX 21-DEC-2000; 2000US-00747835.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-00552317.

XX 20-JUN-2000; 2000US-00598042.

XX 19-JUL-2000; 2000US-00620312.

XX 31-AUG-2000; 2000US-00653450.

XX 04-DEC-2000; 2000US-00729739.

XX (YAMA/) YAMAZAKI V.
PA (TANG/) TANG Y T.

PA (LIUC/) LIU C.
PA (ZHOU/) ZHOU P.
PA (WANG/) WANG D.
PA (ZHANG/) ZHANG J.
PA (ZHANG/) ZHANG J.
PA (REN/) REN F.
PA (ASUN/) ASUNDI V.
PA (DRMA/) DRMANAC R T.
XX
PI Yamazaki V, Tang YT, Liu C, Zhou P, Wang D, Zhang J, Ren F,
PI Asundi V, Drmanac RT;
XX
XX WPI; 2003-174164/17.
XX
PT Isolated polynucleotide for treating, e.g. ophthalmic diseases, comprises
PT specified nucleotide sequences, or their translated protein coding
PT portion, mature protein coding portion, extracellular portion, or active
XX domain.
XX
PS Example 3; Fig 3; 158pp; English.
XX
CC The invention relates to human G protein-coupled receptor-like (GPCR-
CC like) proteins and the polynucleotides encoding them. The isolated
CC sequences are used for the treatment of diseases of ophthalmic,
CC neurological, immunological, or nephritic systems. They may also be used
CC to treat hormonal dysfunction, cancer and other neoplasia,
CC atherosclerosis, diabetes, in treating various anaemias or for use in
CC erythroid precursors and/or erythroid cells. The sequences may be used to
CC promote better or faster closure of non-healing wounds, including
CC pressure ulcers, ulcers associated with vascular insufficiency and
CC surgical and traumatic wounds and in generation or regeneration of other
CC tissues, e.g. organs (e.g. pancreas, liver, intestine, kidney, skin,
CC endothelium), muscle (smooth, skeletal, or cardiac) and vascular tissue
CC or for promoting the growth of cells comprising such tissues. Sequences
CC ABU60884-ABU60895 represent human GPCR-related proteins of the invention
XX
SQ Sequence 267 AA;
XX
Query Match 100.0%; Score 93; DB 6; Length 267;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WIDIKQFWYTPPTF 15
DB 211 WIDIKQFWYTPPTF 225
XX
RESULT 4
ABU60887
ID ABU60887 standard; protein; 267 AA.
XX
AC ABU60887;
XX
DT 08-MAY-2003 (first entry)
XX
DE Human G protein-coupled receptor-related (GPCR-related) protein #4.
XX
KW Human; G protein-coupled receptor; GPCR; receptor; ophthalmic disease;
KW neurological disease; immunological disease; nephritic disease; cancer;
KW hormonal dysfunction; neoplasm; atherosclerosis; diabetes; chemotherapy;
KW anaemia; irradiation; erythroid precursor; erythroid cell; ulcer; organ;
KW pressure ulcer; vascular insufficiency; surgical wound; traumatic wound;
KW pancreas; liver; intestine; kidney; skin; endothelium; skeletal muscle;
KW smooth muscle; cardiac muscle; vascular tissue.
XX
OS Homo sapiens.
XX
EN US2002146692-A1.
XX
PD 10-OCT-2002.
XX
PF 21-DEC-2000; 2000US-00747835.
XX

PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 31-AUG-2000; 2000US-00653450.
PR 04-DEC-2000; 2000US-00729739.
XX
PA (YAMA/) YAMAZAKI V.
PA (TANG/) TANG Y T.
PA (LIUC/) LIU C.
PA (ZHOU/) ZHOU P.
PA (WANG/) WANG D.
PA (ZHANG/) ZHANG J.
PA (REN/) REN F.
PA (ASUN/) ASUNDI V.
PA (DRMA/) DRMANAC R T.
XX
PI Yamazaki V, Tang YT, Liu C, Zhou P, Wang D, Zhang J, Ren F,
PI Asundi V, Drmanac RT;
XX
XX WPI; 2003-174164/17.
XX
PT Isolated polynucleotide for treating, e.g. ophthalmic diseases, comprises
PT specified nucleotide sequences, or their translated protein coding
PT portion, mature protein coding portion, extracellular portion, or active
XX domain.
XX
PS Example 3; Fig 4; 158pp; English.
XX
CC The invention relates to human G protein-coupled receptor-like (GPCR-
CC like) proteins and the polynucleotides encoding them. The isolated
CC sequences are used for the treatment of diseases of ophthalmic,
CC neurological, immunological, or nephritic systems. They may also be used
CC to treat hormonal dysfunction, cancer and other neoplasia,
CC atherosclerosis, diabetes, in treating various anaemias or for use in
CC erythroid precursors and/or erythroid cells. The sequences may be used to
CC promote better or faster closure of non-healing wounds, including
CC pressure ulcers, ulcers associated with vascular insufficiency and
CC surgical and traumatic wounds and in generation or regeneration of other
CC tissues, e.g. organs (e.g. pancreas, liver, intestine, kidney, skin,
CC endothelium), muscle (smooth, skeletal, or cardiac) and vascular tissue
CC or for promoting the growth of cells comprising such tissues. Sequences
CC ABU60884-ABU60895 represent human GPCR-related proteins of the invention
XX
SQ Sequence 267 AA;
XX
Query Match 100.0%; Score 93; DB 6; Length 267;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WIDIKQFWYTPPTF 15
DB 211 WIDIKQFWYTPPTF 225
XX
RESULT 5
AAV58194
ID AAV58194 standard; protein; 339 AA.
XX
AC AAV58194;
XX
DT 14-MAR-2000 (first entry)
XX
DE Human STRAP-1 protein.
XX
KW Serpentine transmembrane antigen of the prostate; STRAP-1; prostate;
KW transmembrane domain; type IIIa membrane protein; expression; cancer;
KW prostate cancer; bladder cancer; colon cancer; pancreatic cancer;
KW ovarian cancer; tumour antigen; immunisation; immune response; cellular;
KW humoral; anticancer vaccine; antibody; detection; diagnosis; prognosis;
KW monitoring; susceptibility; therapeutic inhibitor; drug targeting;
KW recombinant protein.
XX

Accession	Protein Name	Location/Qualifiers
XX	Homo sapiens.	
XX		
XX	Key	Location/Qualifiers
FT	Region	1..69
FT		/note= "intracellular region 1"
FT	Domain	70..91
FT		/note= "Transmembrane domain 1"
FT	Region	92..113
FT		/note= "Extracellular region 1 (AAV58198) "
FT	Domain	114..136
FT		/note= "Transmembrane domain 2"
FT	Region	137..162
FT		/note= "intracellular region 2"
FT	Domain	163..184
FT		/note= "Transmembrane domain 3"
FT	Region	185..218
FT		/note= "Extracellular region 2 (AAV58199) "
FT	Domain	219..241
FT		/note= "Transmembrane domain 4"
FT	Region	242..253
FT		/note= "intracellular region 3"
FT	Domain	252..276
FT		/note= "Transmembrane domain 5"
FT	Region	277..291
FT		/note= "Extracellular region 3 (AAV58200) "
FT	Domain	292..313
FT		/note= "Transmembrane domain 6"
FT	Region	314..339
FT		/note= "intracellular region 4"
XX		
XX	WC0962941-A2.	
XX		
XX	09-DEC-1999.	
XX		
PF	01-UTN-1999;	99WC-US012157.
XX		
PR	01-UTN-1998;	98US-0087520P.
XX		
PR	30-UTN-1998;	98US-0091183P.
XX		
PA	(UROG-) UROGENESYS INC.	
PA	(AFAR/) AFAR D E.	
PA	(HUBE/) HUBERT R S.	
PA	(LEON/) LEONG K.	
PA	(RAIT/) RAITANO A B.	
PA	(SAFF/) SAFFRAN D C.	
XX		
PI	Afar DE, Hubert RS, leong K, Raitano AB, Safran DC;	
DR	WPI, 2000-072832/06.	
DR	N-PSDB; AAZ49395, AAZ49396.	
XX		
PT	Novel proteins useful as diagnostic markers and therapeutic targets,	
PT	particularly for prostatic cancer.	
XX		
PS	Claim 1; Fig 1A; 83pp; English.	
XX		
CC	This sequence represents a novel human protein, STRAP-1 (serpentine	
CC	transmembrane antigen of the prostate). STRAP-1 is the prototype member	
CC	of the STRAP family of proteins (AAV58194-Y58197) which exhibit a high	
CC	degree of structural conservation, but which show no significant	
CC	structural homology to known human proteins. The STRAP-1 gene has been	
CC	localised to chromosome 7p22. STRAP-1 is thought to be a type IIa	
CC	membrane protein and is expressed predominantly in prostate cells in	
CC	normal human tissues. Structurally, STRAP-1 is a 339 amino acid protein	
CC	characterised by six transmembrane domains and intracellular N- and C-	
CC	termini, suggesting that it folds in a "serpentine" manner into three	
CC	extracellular and two intracellular loops. STRAP-1 mRNA and protein	
CC	expression is maintained at high levels and throughout all stages of	
CC	prostate cancer. STRAP-1 mRNA and/or protein is also overexpressed in	
CC	certain other cancers, including bladder, colon, pancreatic and ovarian	
CC	cancer. The function of the STRAP proteins is not known. They may be ion	
CC	channels (from the presence of six transmembrane domains, a feature which	

is shared by certain ion channels) or gap-junction proteins (from immunohistochemical staining). STRAP-1 and STRAP-2 are cell surface and tumour antigens. Immunisation with a STRAP protein induces cellular and humoral immune responses against STRAP-expressing cells. STRAP proteins may be used to identify specific-binding agents, to produce anticancer vaccines and to generate specific antibodies. The antibodies may be used for detection, prognosis, and monitoring of cancers (or susceptibility to cancer), as therapeutic inhibitors or to target therapeutic agents to their site of action. STRAP nucleic acids may be used for recombinant protein production, as diagnostic and prognostic reagents, for identifying STRAP-expressing cells for screening inhibitors of STRAP expression and for therapeutic modulation/inhibition of STRAP expression. Since high levels of STRAP proteins are exposed on the cell surface, they are easily targeted by systemically administered agents, and because they are expressed mainly on prostatic epithelial cells, agents targeted to them should have minimal side effects on other tissues

Sequence 339 AA:

	Query Match	100.0%;	Score 93;	DB 3;	Length 339;
	Best Local Similarity	100.0%;	Pred. No. 1.3e-05;		
	Matches 15;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Dy	1 WDIKQFWYTPPTF 15 				
Dd	277 WDIKQFWYTPPTF 291				

RESULT 6
AA001282
ID AA001282 standard; protein; 339 AA.
XX AA001282;
XX
DT 04-OCT-2001 (first entry)
DE P789P amino acid sequence.
XX
XX Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KM cytotatic; gene therapy; metastasis.
OS Homo sapiens.
PN WO200151633-A2.
PD 19-JUL-2001.
XX
XX 16-JAN-2001; 2001WO-US001574.
PF 14-JAN-2000; 2000US-00483672.
PR XX
PA (CORI-) CORIXA CORP.
PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI Kalos MD, Fanger GR, Day CH, Rector MM, Stolk JA, Skeiky YAM,
PI Wang A, Meagher MJ;
XX
DR WPI; 2001-425873/45.

New polynucleotide encoding a prostate-specific protein, for diagnosing, monitoring and treating prostate cancer in a patient and for use in vaccines.

Claim 2; Page 510-512; 543pp; English.

The present invention describes polynucleotide sequences (I) which encode prostate-specific proteins (II). (I) and (II) have cytostatic activity, and can be used in vaccine production and gene therapy. (I), (II), T cells antibodies to (II), fusion proteins comprising (II), and isolated T cells prepared using (I) or (II) are used treat cancer in a patient. (I) and the antibodies are also used in the detection of cancer in a patient. The cancer that is diagnosed or treated is particularly prostate cancer. (I) and (II) can be used in vaccines. The antibodies or (I) can be used for

CC monitoring the progression of cancer in a patient. (I) and (II) can also
CC be used to improve diagnostic and therapeutic methods for prostate
CC cancer. They can indicate the level of metastasis as well as the prostate
CC volume. AAH93357 to AAH93944 and AAM0115 to AAM0318 represent
CC polynucleotide and amino acid sequences used in the exemplification of
CC the present invention
XX

Sequence 339 AA:

Query Match 100.0%; Score 93; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 WIDIKQFWYTPPTF 15
277 WIDIKQFWYTPPTF 291

RESULT 7

AAU69927 standard; protein; 339 AA.

AAU69927;

30-JAN-2002 (first entry)

Human prostate cDNA encoded protein #72.

Human; prostate cancer; cytostatic; immunostimulant; tumour; immunogen.

Homo sapiens.

WO200173032-A2.

04-OCT-2001.

27-MAR-2001; 2001WO-US009919.

27-MAR-2000; 2000US-00536857.

09-MAY-2000; 2000US-00568100.

12-MAY-2000; 2000US-00570737.

13-JUN-2000; 2000US-00593793.

27-JUN-2000; 2000US-00605783.

09-AUG-2000; 2000US-00636215.

29-AUG-2000; 2000US-00651236.

06-SEP-2000; 2000US-00657279.

10-OCT-2000; 2000US-00679426.

10-OCT-2000; 2000US-00685166.

09-NOV-2000; 2000US-00709729.

(CORI-) CORIYA CORP.

Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD,
Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
Li SX, Wang A, Skeiky YAM, Hepler WT, Henderson RA;
WPI; 2001-639232/73.
N-PSDB; AAS64160.

New human prostate-specific polypeptides and polynucleotides useful for
the diagnosis and treatment of cancer, especially prostate cancer.

Claim 2; Page 549; 579pp; English.

The invention relates to isolated prostate-specific polynucleotides,
polypeptides, fusion proteins of the polypeptides, antibodies raised
against the polypeptides (or antigenic epitopes derived from them) and
antigen-presenting cells expressing the polypeptides. The antibodies are
useful for detecting the presence of cancer, especially prostate cancer.
The polypeptides, polynucleotides and the antigen-presenting cells are
useful for stimulating and/or expanding T cells specific for a tumour
protein, and for inhibiting the development of cancer especially prostate
cancer. Compositions comprising the polynucleotide and/or polypeptide are

CC useful for stimulating an immune response, and for treating cancer. The
CC oligonucleotide is useful for detecting cancer. The present sequence is a
CC prostate specific polypeptide of the invention
XX

Sequence 339 AA:

Query Match 100.0%; Score 93; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 WIDIKQFWYTPPTF 15
277 WIDIKQFWYTPPTF 291

RESULT 8

AAM78845 standard; protein; 339 AA.

AAM78845;

06-NOV-2001 (first entry)

Human protein SEQ ID NO 1507.

Human; cytokine; cell proliferation; cell differentiation; gene therapy;
vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

tissue growth factor; immunomodulatory; cancer; leukaemia;

nervous system disorder; arthritis; inflammation.

Homo sapiens.

WO200157190-A2.

09-AUG-2001.

05-FEB-2001; 2001WO-US004098.

03-FEB-2000; 2000US-00496914.

27-APR-2000; 2000US-00560875.

20-JUN-2000; 2000US-00598075.

19-JUL-2000; 2000US-00620325.

01-SEP-2000; 2000US-0064936.

15-SEP-2000; 2000US-00663561.

20-OCT-2000; 2000US-00693325.

30-NOV-2000; 2000US-00728422.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y,
Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
Xue AJ, Yang Y, Wejhrman T, Goodrich R;
WPI; 2001-476283/51.
N-PSDB; AAK51978.

Nucleic acids encoding polypeptides with cytokine-like activities, useful
in diagnosis and gene therapy.

Claim 20; Page 3800-3801; 6221pp; English.

The invention relates to polynucleotides (AAK51456-AAK53435) and the
encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to
cytokine, cell proliferation or cell differentiation or which may induce
production of other cytokines in other cell populations. The
polynucleotides and polypeptides are useful in gene therapy, vaccines or
peptide therapy. The polypeptides have various cytokine-like activities,
e.g. stem cell growth factor activity, haematopoiesis regulating
activity, tissue growth factor activity, immunomodulatory activity and
activin/inhibin activity and may be useful in the diagnosis and/or
treatment of cancer, leukaemia, nervous system disorders, arthritis and
inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
(AAK52582) and 3666 (AAM8020) are omitted as the relevant pages from the

```
CC sequence listing were missing at the time of publication
XX
SQ Sequence 339 AA;
Query Match 100.0%; Score 93; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 WIDIKQFWYTPPTF 15
DB 277 WIDIKQFWYTPPTF 291
RESULT 9
ABU71818
ID ABU71818 standard; protein; 339 AA.
XX
AC ABU71818;
XX
DT 10-JUN-2003 (first entry)
XX
DE Prostate cancer associated protein #53.
XX
KW Prostate cancer; vaccine; gene therapy; cytostatic; fusion protein;
KW immunogen; cancer; prostate specific antigen; PSA;
KW prostatic acid phosphatase; PAP; prostate specific membrane antigen;
KW PSMA.
XX
OS Homo sapiens.
XX
PN US2002192763-A1.
XX
PD 19-DEC-2002.
XX
PF 29-JUN-2001; 2001US-00895793.
XX
PR 04-OCT-1999; 99US-0157455P.
PR 04-OCT-2000; 2000US-00679272.
PR 28-MAR-2001; 2001US-00822827.
XX
(XUUU/) XU J.
PA (DILL/) DILLON D C.
PA (MITC/) MITCHAM J L.
PA (HARL/) HARLOCKER S L.
PA (JIANG/) JIANG Y.
PA (KALO/) KALOS M D.
PA (FRANG/) FRANGER G R.
PA (REIT/) REITTER M W.
PA (STOL/) STOLK J A.
PA (DAYC/) DAY C H.
PA (VEDV/) VEDVICK T S.
PA (CART/) CARTER D.
PA (LISX/) LI S X.
PA (WANG/) WANG A.
PA (SKEI/) SKEIKY Y A W.
PA (HEPL/) HEPLER W T.
PA (HEND/) HENDERSON R A.
PA (HURA/) HURAL J.
PA (MCNE/) MCNETIL P D.
PA (HOUN/) Houghton R L.
PA (DBAS/) Y DE BASSOLS C V.
PA (FOYT/) FOY T M.
XX
XU J, Dillon DC, Mitcham JL, Harlocker SH, Jiang Y, Kalos MD,
PI Fanger GR, Reitter MW, Stolk JA, Day CH, Vedvick TS, Carter D,
PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA, Hural J,
PI McNeill PD, Houghton RL, Y De Bassolecv, Foy TM;
XX
DR WPI; 2001-245062/25.
DR N-PSDB; ACA59965.
XX
PT Prostate specific protein and its encoding polynucleotide, useful for the
PT treatment and diagnosis of prostate cancer.
```

```
XX
PS Example 5; SEQ ID NO 879; 85pp; English.
XX
CC The invention describes a fusion protein comprising at least one amino
CC acid sequence of immunogenic portions of any of the 3 sequences not
CC defined in the specification, or sequences having at least 70 or 90 %
CC sequence identity to any one of the 35 sequences defined in the USPTO web
CC site, which is encoded by any of the 4 nucleotide sequences not defined
CC in the specification. The fusion protein, composition and methods are
CC useful for diagnosing, preventing and/or treating cancer, particularly
CC prostate cancer. The proteins are useful as markers to indicate the
CC presence or absence of cancer. This is the amino acid sequence of a
CC prostate cancer therapy associated protein. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?DocID=US20020192763
XX
SQ Sequence 339 AA;
Query Match 100.0%; Score 93; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 WIDIKQFWYTPPTF 15
DB 277 WIDIKQFWYTPPTF 291
RESULT 10
ABB95387
ID ABB95387 standard; protein; 339 AA.
XX
AC ABB95387;
XX
DT 19-JUL-2002 (first entry)
XX
DE Human P789P protein SEQ ID NO 879.
XX
KW Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
KW gene therapy.
XX
OS Homo sapiens.
XX
PN US2002022248-A1.
XX
PD 21-FEB-2002.
XX
PF 12-JAN-2001; 2001US-00759143.
XX
PR 25-FEB-1997; 97US-00806099.
PR 01-AUG-1997; 97US-00904804.
PR 10-FEB-1998; 98US-00020956.
PR 25-FEB-1998; 98US-00030607.
PR 14-JUL-1998; 98US-00115453.
PR 23-SEP-1998; 98US-00159812.
PR 15-JAN-1999; 99US-00232149.
PR 09-APR-1999; 99US-00288946.
PR 13-JUL-1999; 99US-00352616.
PR 12-NOV-1999; 99US-00439313.
PR 18-NOV-1999; 99US-00443586.
PR 14-JAN-2000; 2000US-00483672.
PR 27-MAR-2000; 2000US-00536857.
PR 09-MAY-2000; 2000US-00568100.
PR 12-MAY-2000; 2000US-00570737.
PR 13-JUN-2000; 2000US-00593793.
PR 27-JUN-2000; 2000US-00605783.
PR 10-AUG-2000; 2000US-00636215.
PR 29-AUG-2000; 2000US-00651236.
PR 06-SEP-2000; 2000US-00657279.
PR 02-OCT-2000; 2000US-00679426.
PR 10-OCT-2000; 2000US-00685166.
XX
(XUUU/) XU J.
```


DR WPI; 2003-313240/30.
DR N-PSDB; ACD02597.
XX
PT New composition comprising a substance that modulates the status of a
PT STEAP-1-related protein, useful for treating and detecting cancer.
XX
PS Example 2; Fig 2A; 248bp; English.
XX
CC The invention describes a composition comprising a substance that
CC modulates the status of a protein (I) of 340 or 283 amino acids, or of
CC any of the 15 sequences of 259 amino acids, given in the specification,
CC or a molecule that is modulated by the protein, where the status of the
CC cell that expresses the protein is modulated. The compositions, proteins,
CC polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
XX
SQ Sequence 339 AA;
XX
Query Match 100.0%; Score 93; DB 6; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 WIDIKQFWYTPPTF 15
Db 277 WIDIKQFWYTPPTF 291
XX
RESULT 13
ID ABU98414 standard; protein; 339 AA.
AC ABU98414;
XX
XX 31-JUL-2003 (first entry)
DT
XX
DE STEAP-1 variant 1.
XX
XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
XX cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX vaccine.
XX
OS Homo sapiens.
XX
XX W02003022995-A2.
XX
XX 20-MAR-2003.
PD
XX
PF 06-SEP-2002; 2002WO-US028371.
XX
XX 06-SEP-2001; 2001US-0317840P.
XX
XX 05-APR-2002; 2002US-0370387P.
XX
XX (AGEN-) AGENSYS INC.
XX
XX Farris M, Ge W, Raitano AB, Challita-Rid PM, Jakobovits A;
XX
XX WPI; 2003-313240/30.
XX
XX New composition comprising a substance that modulates the status of a
XX STEAP-1-related protein, useful for treating and detecting cancer.
XX
XX Example 53; Page 173; 248bp; English.
XX
XX The invention describes a composition comprising a substance that
XX modulates the status of a protein (I) of 340 or 283 amino acids, or of
XX any of the 15 sequences of 259 amino acids, given in the specification,
XX or a molecule that is modulated by the protein, where the status of the
XX cell that expresses the protein is modulated. The compositions, proteins,
XX

CC polynucleotides and methods are useful for treating and detecting cancer.
CC The STEAP-1-related proteins are useful for generating cancer vaccines.
CC The polynucleotides are useful as tools for delineating, with greater
CC precision, cytogenetic abnormalities in the chromosomal region that
CC encodes STEAP-1 that may contribute to the malignant phenotype. This is
CC the amino acid sequence of a variant of human six transmembrane
CC epithelial antigen of the prostate or STEAP-1
XX
SQ Sequence 339 AA;
XX
Query Match 100.0%; Score 93; DB 6; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 WIDIKQFWYTPPTF 15
Db 277 WIDIKQFWYTPPTF 291
XX
RESULT 14
ID ABU98430 standard; protein; 339 AA.
AC ABU98430;
XX
XX 31-JUL-2003 (first entry)
DT
XX
DE STEAP-1 variant BP1D4 v.1 #1.
XX
XX STEAP-1; six transmembrane epithelial antigen of the prostate; cancer;
XX cancer vaccine; delineation; cytogenetic abnormality; cytostatic;
XX vaccine.
XX
OS Homo sapiens.
XX
XX W02003022995-A2.
XX
XX 20-MAR-2003.
PD
XX
PF 06-SEP-2002; 2002WO-US028371.
XX
XX 06-SEP-2001; 2001US-0317840P.
XX
XX 05-APR-2002; 2002US-0370387P.
XX
XX (AGEN-) AGENSYS INC.
XX
XX Farris M, Ge W, Raitano AB, Challita-Rid PM, Jakobovits A;
XX
XX WPI; 2003-313240/30.
XX
XX New composition comprising a substance that modulates the status of a
XX STEAP-1-related protein, useful for treating and detecting cancer.
XX
XX Example 9; Fig 13A; 248bp; English.
XX
XX The invention describes a composition comprising a substance that
XX modulates the status of a protein (I) of 340 or 283 amino acids, or of
XX any of the 15 sequences of 259 amino acids, given in the specification,
XX or a molecule that is modulated by the protein, where the status of the
XX cell that expresses the protein is modulated. The compositions, proteins,
XX polynucleotides and methods are useful for treating and detecting cancer.
XX The STEAP-1-related proteins are useful for generating cancer vaccines.
XX The polynucleotides are useful as tools for delineating, with greater
XX precision, cytogenetic abnormalities in the chromosomal region that
XX encodes STEAP-1 that may contribute to the malignant phenotype. This is
XX the amino acid sequence of a variant of human six transmembrane
XX epithelial antigen of the prostate or STEAP-1
XX
SQ Sequence 339 AA;
XX
Query Match 100.0%; Score 93; DB 6; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WIDIKQFVWYTPPTF 15
 |||||
 Db 277 WIDIKQFVWYTPPTF 291

RESULT 15

ABR54499

ID ABR54499 standard; protein; 339 AA.

AC ABR54499;

DT 28-AUG-2003 (first entry)

DE Prostate tumour specific protein SEQ ID 879.

KW Cytostatic; gene therapy; prostate-specific protein; PSP; human;
 immune response; prostate cancer.

OS Homo sapiens.

PN WO200289747-A2.

PD 14-NOV-2002.

PF 09-MAY-2002; 2002WO-US014753.

PR 09-MAY-2001; 2001US-00852911.

PR 29-JUN-2001; 2001US-00895814.

PR 10-DEC-2001; 2001US-00012896.

(CORI-) CORIXA CORP.

PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Henderson RA,
 PI Kalos MD, Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS,
 PI Carter D, Li SX, Wang A, Skeiky YAM, Hepler WT, Hurai J,
 PI McNeill PD, Houghton KU, Vinals Y De BassolsC, Foy TM, Watanabe Y,
 PI Deng T;

DR WPI; 2003-167130/16.

PT New prostate-specific proteins and genes, useful in gene therapy,
 particularly for stimulating an immune response in a patient, or treating
 prostate cancer in a patient, as well as for diagnosing prostate cancer
 in a patient.

PS Example 5; Page 618-619; 691pp; English.

CC The present invention relates to novel prostate-specific proteins (PSP)
 and their coding sequences. The PSPs and their coding sequences are
 useful for stimulating an immune response in a patient, or for treating
 prostate cancer in a patient and for determining, detecting or diagnosing
 the presence of a cancer in a patient. The present sequence was used to
 illustrate the invention

SQ Sequence 339 AA;

Query Match 100.0%; Score 93; DB 6; Length 339;
 Best Local Similarity 100.0%; Pred. NO. 1.3e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WIDIKQFVWYTPPTF 15
 |||||
 Db 277 WIDIKQFVWYTPPTF 291

Search completed: July 12, 2004, 14:07:24
 Job time : 33.9651 secs

This Page Blank (uspto)

This Page Blank (uspto)